



## **Editor in Chief**

Jürgen Rudolph, *Kaplan Higher Education Academy, Singapore*

## **Senior Editor**

Mohamed Fadhil bin Mohamed Ismail, *Kaplan Higher Education Academy, Singapore*

## **Chairperson**

Joseph Crawford, *University of Tasmania*

## **Deputy Chairperson**

Tania Aspland, *Kaplan Business School Australia*

## **Associate Editors**

Michael Agyemang Adarkwah, *Friedrich Schiller University Jena, Germany*

Sayan Dey, *Alliance University, Bangalore*

Matthias Glowatz, *University College Dublin, Ireland*

James Kwan, *Embry-Riddle Aeronautical University, Singapore*

Lydia Lymperis, *National Institute of Teaching (NlOT), United Kingdom*

Kyriaki Koukouraki, *King's College London, England*

Tarin Ong, *Murdoch University, Singapore*

Caleb Or, *Singapore Institute of Technology, Singapore*

Valentine Joseph Owan, *Ultimate Research Network, Nigeria*

Mike Perkins, *British University Vietnam, Vietnam*

Faiza Qureshi, *Murdoch University, Dubai*

Felix Sanni, *Fescosof Data Solutions, Nigeria*

Jennifer Skipp, *Trier University, Germany*

Vanessa Stafford, *Kaplan Business School, Australia*

Effie Steriopoulos, *William Angliss Institute, Australia*

Fiona Tang, *Kaplan Business School, Australia*

## **Journal Manager**

Shannon Tan, *Kaplan Higher Education Academy, Singapore*

## **Event Manager**

Pauline Seah, *Kaplan Higher Education Academy, Singapore*

## Editorial Board

- James Adonopoulos, *Kaplan Business School, Australia*  
Khairudin Aljunied, *National University of Singapore*  
Martin Andrew, *Otago Polytechnic, New Zealand*  
Abhishek Bhati, *James Cook University, Singapore*  
Begüm Burak, *Independent Researcher, Turkey*  
Rob Burton, *Northumbria University, UK*  
Kerryn Butler-Henderson, *Charles Sturt University, Australia*  
Ailson De Moraes, *Royal Holloway, University of London, UK*  
Fotini Diamantidaki, *University College London, UK*  
Lena Itangata, *University of Portsmouth, UK*  
Margarita Kefalaki, *Communication Institute of Greece*  
Sophia Lam, *The Education University of Hong Kong, Hong Kong*  
Bashar Malkawi, *Government of Dubai, United Arab Emirates*  
Jane Murray, *Northumbria University, UK*  
Florence Mwagwabi, *Murdoch University, Singapore*  
Justin O'Brien, *University of Surrey, UK*  
Orna O'Brien, *University College Dublin, Ireland*  
Can-Seng Ooi, *University of Tasmania, Australia*  
Michelle Picard, *Flinders University, Australia*  
Stefan Popenici, *Charles Darwin University, Australia*  
Bina Rai, *National University of Singapore*  
Hasan Saliu, *AAB College, Kosovo*  
Sam Choon-Yin, *Kaplan Higher Education Academy, Singapore*  
Arthur Shelley, *RMIT University, Melbourne, Australia*  
Becky Shelley, *University of Tasmania, Australia*  
Stephen Shukaitis, *University of Essex, UK*  
Eunice Tan, *Singapore Institute of Management, Singapore*  
Renée Tan, *Institute for Adult Learning, Singapore*  
Samson Tan, *Institute for Adult Learning, Singapore*  
Peter Waring, *Murdoch University, Australia*

---

# Contents

---

## Editorial

Joyce's *Odyssey* in *Ulysses*. A celebration of human ingenuity and an indictment of the terminal dullness of generative AI 07-21

Jürgen Rudolph, Shannon Tan & Mohamed Fadhil bin Mohamed Ismail

---

Ethical concerns for using artificial intelligence chatbots in research and publication: Evidences from Saudi Arabia 93-103

Abu Elnasr Sobaih

---

Generative Artificial Intelligence in distance education: Transformations, challenges, and impact on academic integrity and student voice 104-114

Kershnee Sevnanarayan, & Mary-Anne Potter

---

Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review 115-126

Chaka Chaka

---

Accuracy pecking order – How 30 AI detectors stack up in detecting generative artificial intelligence content in university English L1 and English L2 student essays 127-139

Chaka Chaka

---

Artificial intelligence in higher education database (AIHE V1): Introducing an open-access repository 140-148

Mohamed Fadhil bin Mohamed Ismail, Joseph Crawford, Shannon Tan, Jürgen Rudolph, Eunice Tan, Pauline Seah, Fiona Xiaofei Tang, Florence Ng, Laura Visser Kaldenbach, Anand Naidu, Vanessa Stafford, & Michelle Kane

---

Digital versus classroom discussions: Motivation and self-efficacy outcomes in speaking courses via Gather.town 149-161

Nimer Abusalim, Mohammad Rayyan, Sarah Alshanmy, Sharif Alghazo, & Ghaleb Rababah

---

Utilizing head simulation training in dental school education: Time and cost implications 162-167

Samuel E. Inkabi, Isaac T. Inkabi, Shirley Inkabi, Adu Baffour, Adwoa A. Asubonteng, Lydia E. Boamah, David Benefo, Philip B. Frimpong, Obed A. Asamoah, & Michael Oppong-Atuahene

---

Self-perception of university teachers on their digital teaching competence: The case of Peru 168-181

Mercedes Rojas-Osorio, Shyla Del-Aguila-Arcental-es, & Aldo Alvarez-Risco

---

## Commentary

Artificial intelligence and graduate employability: What should we teach Generation AI? 22-25

Peter Waring

---

## Research articles

Thematic exploration of educational research after the COVID pandemic through topic modelling 26-34

Indu Bala & Lewis Mitchell

---

Is ChatGPT an opportunity or a threat? Preventive strategies employed by academics related to a GenAI-based LLM at a faculty of education 35-45

Micheal M Van Wyk

---

Higher education assessment practice in the era of generative AI tools 45-56

Bayode Ogunleye, Kudirat Ibilola Zakariyyah, Oluwaseun Ajao, Olakunle Olayinka, & Hemlata Sharma

---

Awareness, benefits, threats, attitudes, and satisfaction with AI tools among Asian and African higher education staff and students 57-64

Muayyad Ahmad, Maha Subih, Mirna Fawaz, Hanan Alnuqaidan, Ashraf Abuejheisheh, Vian Naqshbandi, & Fadwa Alhalaqa

---

Exploring graduate students' perception and adoption of AI chatbots in Zimbabwe: Balancing pedagogical innovation and development of higher-order cognitive skills 65-75

Arthur William Fodouop Kouam, & Regis Misheal Muchowe

---

Artificial Intelligence (AI) in academic research. A multi-group analysis of students' awareness and perceptions using gender and programme type 76-92

Ofem Usani Joseph, Iyam Mary Arikpo, Ovat Sylvia Victor, Nwogwugwu Chidirim, Anake Paulina Mbua, Udeh Maryrose Ify, & Otu Bernard Diwa

---

---

# Contents

---

'Just get them over the line': Neoliberalism and the execution of 'excellence'	182-192	Do graduate courses in a HyFlex mode foster emotional, cognitive and behavioral engagement? A consideration	278-290
Martin Benedict Andrew		David Santandreu Calonge, Mariam Aman Shah, Melissa Connor, Patrik Hultberg, & Pablo Medina Aguerrebere	
The Great Resignation: The simple joys of not belonging	193-205	Student engagement and fostering ownership of learning	291-302
Martin Benedict Andrew		Ajay Khatter, Krishna Thalaachawr, & Mark Blyth	
Neoliberalism and the violation of students' rights: The case of English language education	206-213	Evaluation of a research training workshop for academic staff in tertiary institutions: A Kirkpatrick model approach	303-313
Ehsan Nikpouya, & Gholamreza Zareian		Caroline Ochuko Alordiah	
Metrics in research impact assessment and grant funding: Insights from researchers in the "Reviewer 2 Must Be Stopped!" Facebook group	214-225	Professional development activities of English language lecturers in Vietnam through the lens of sociocultural theory	314-324
Valentine J. Owan, Victor U. Agama, John O. Odey, & Delight O. Idika		Le Pham Huong Hoai, Tran Thi Thuong Thanh, & Hung Phu Bui	
A content analysis of tweets on toxic doctoral supervision	226-234	Development and validation of an instrument to measure expectancy for success and subjective task value constructs in the context of higher education	325-335
Oluchi Ojinnamma Okere		Wan Hoong Wong, & Elaine Chapman	
The challenge of making relationships central in online cultural safety education	235-246	Advice from retired secondary school principals in Ireland on how to lead as a principal	336-345
Robyn Moore, Kim McLeod, & Natalia Maystorovich Chulio		Finn Ó Murchú, & Tom O'Donoghue	
A conceptual, strategic and implementation framework for the Scholarship of Learning and Teaching	247-255	<b>Interviews</b>	
Melanie Williams		Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin	346-363
Micro-credentials in higher and vocational education: An innovation or a disruption? A review of the literature	256-265	Rose Luckin, Jürgen Rudolph, Martin Grünert, & Shannon Tan	
Jeremy Hanshaw		'Failing well' in teaching about race, racism and white supremacy. An interview with Stephen Brookfield	364-389
Organisational resilience in a higher education institution: Maintaining academic continuity, academic rigour and student experience in the face of major disruption (Covid-19 pandemic)	266-277	Stephen D. Brookfield, Jürgen Rudolph, & Shannon Tan	
Adeela Ahmed Shafi, & Tristan Middleton			

---

# Contents

---

## Ed-tech review

---

The use of Generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT 390-395

Mike Perkins, & Jasper Roe

---

## Opinion pieces

Reading between the lines: The necessity of books 396-405

Sam Choon Yin

---

Cargo cultism and the whiteness syndrome: Fake internationalization of private universities of India 406-412

Sayan Dey, & Aritra Chakraborty

---

Rise of the robots: What It means for educators 413-420

Augustine Osamor Ifelebuegu

---

## Brief article

Recognition of foreign professional degrees in Peru: Processes and strategies for improvement 421-426

Gianfranco Cavagnari-Bruce, Shyla Del-Aguila-Arcenales, & Vladimir Rodriguez Cairo

---

## Book reviews

Book review. Brookfield, Stephen D., Rudolph, Jürgen & Tan, Shannon (2024). Teaching well: Understanding key dynamics of learning-centre classrooms. Routledge. 427-428

Peter Waring

---

Book review of Lindgren, Simon (Ed., 2023). Handbook of critical studies of artificial intelligence. Edward Elgar. 429-435

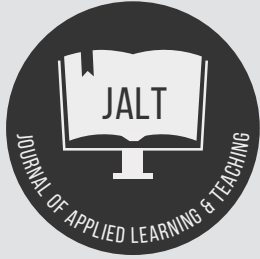
Jürgen Rudolph

---

Book review of Benedict du Boulay, Antonija Mitrovic, & Kalina Yacef (Eds., 2023). Handbook of artificial intelligence in education. Edward Elgar. 436-438

Jürgen Rudolph

---



---

## Joyce's *Odyssey*. A celebration of human ingenuity in *Ulysses* and an indictment of the mediocrity of generative AI

---

Jürgen Rudolph<sup>A</sup>

A Director of Research &amp; Learning Innovation, Kaplan Higher Education Academy

Shannon Tan<sup>B</sup>

B Research Executive, Kaplan Higher Education Academy

Fadhil Mohamed bin

C Head of Academic Support and Curriculum Development, Kaplan Higher Education Academy

Mohamed Ismail<sup>C</sup>DOI: <https://doi.org/10.37074/jalt.2024.7.1.1>

---

### Abstract

This editorial reflects on James Joyce's modernist novel *Ulysses*, first published 100 years ago in 1922. We reconstruct *Ulysses*'s revolutionary redefinition of the novel genre, its critical reception, and the immense challenges Joyce faced in writing, printing, and publishing the work. Narrating the genesis of *Ulysses* is a celebration of human ingenuity and perseverance in the face of daunting obstacles. We contrast Joyce's brilliant literary achievement with the comparatively inferior outputs of much-hyped generative artificial intelligence (GenAI) chatbots. We discuss how the excessive caution and censorship exhibited by leading generative AI systems like ChatGPT undermine the free exchange of ideas, in stark contrast to the liberation of expression embodied by *Ulysses*. Finally, we consider the implications of these insights for effective teaching practices and visionary leadership in higher education, emphasising the vital role of cultivating broad intellectual engagement and critical thinking skills among students and faculty. Our editorial also provides an overview of the many human-created gems in our latest journal issue.

**Keywords:** AI; artificial intelligence; fiction; generative AI; generative artificial intelligence; higher education; Homer; human ingenuity; James Joyce; literature; *The Odyssey*; *Ulysses*.

### Introduction to *Ulysses*

Oh rocks. Tell us in plain words.  
(Molly Bloom in Joyce, 2000a, p. 77)

It is ironic that reflecting on artificial intelligence and higher education for the last year has led us – the authors of this Editorial – to increasingly appreciate human intelligence. In this Editorial, we have decided to do something slightly

unusual: to celebrate human literary ingenuity that favourably compares to the terminally dull texts that generative AI tends to spew out (Rudolph et al., 2023a, 2023b). To showcase the superiority of the human intellect, we could have provided innumerable examples from literature, visual art, theatre, film, or music. Instead, we shine our torchlight on only one monumental novel that was published some 100 years ago (in 1922): James Joyce's *Ulysses*. Homer's magnificent work, *The Odyssey*, approximately 2,700 years old and consisting of more than 12,000 lines of hexameter verse (Knox, 2006), serves as a crucial point of reference for Joyce's *Ulysses* and would have been another worthy case in point.

This is not the first instance of discussing a literary masterpiece within a JALT Editorial – Rudolph et al. (2022) discuss *Faust* (Goethe, 1997, 2003) in the context of the Faustian pacts that we enter in the context of neoliberal higher education. Goethe's *Faust* served as an allegory for the Faustian bargains of modernity, notably our unyielding faith in never-ending progress, which has precipitated environmental degradation in the Anthropocene – the epoch we currently live in. We also reflected on how the existential and epistemological crises engendered by the pandemic mirrored Faust's despair over the limitations of his knowledge and the quest for meaning in his teachings. We concluded that *Faust* challenges us to reflect on the essence of human striving and the possibility of redemption, advocating for a critically-tempered hope in the face of adversity and injustice (Rudolph et al., 2022).

While the recent pandemic's challenges framed our exploration of *Faust*, our examination of *Ulysses* navigates through the prevailing generative AI epidemic, characterised by viral hype and hysteria. In the realm of higher education, the erosion of extensive reading habits among students – and, to a lesser extent, educators – signals a troubling trend. This decline diminishes the potential for developing robust writing skills, given the well-established notion that avid readers frequently become more adept writers (Pinker, 2014). Furthermore, the escalating specialisation

within academic disciplines threatens to stifle the breadth of interdisciplinary knowledge epitomised by geniuses of bygone eras like Leonardo da Vinci and Goethe. Such broad-ranging scholarship, which fosters a culture of intellectual curiosity and cross-disciplinary learning, is at risk of being eclipsed by a narrow emphasis on domain-specific expertise. James Joyce's *Ulysses*, a paragon of literary complexity and innovation, serves as a poignant reminder of the richness that broad and deep engagement with literature can offer. We underscore the irreplaceable value of human creativity and intellectual depth by juxtaposing Joyce's opus magnum with the outputs of generative AI (GenAI). While GenAI may offer impressive feats of content generation at breakneck speed, it falls markedly short of replicating the creative syntheses that human artists have achieved through the ages.

This observation is not merely an academic point but has profound implications for good teaching and higher education leadership (Brookfield et al., 2023; Tan et al., 2024). Fostering a culture of wide-ranging intellectual exploration among students and faculty is not just beneficial but essential. Encouraging engagement with works of complex literature like *Ulysses* can serve as a powerful antidote to the narrowing of academic focus, enriching students' educational experiences and equipping them with the creative and critical thinking skills necessary for thriving in a VUCA (volatile, uncertain, complex, and ambiguous) environment. Moreover, for educators and leaders in higher education, championing the value of broad and interdisciplinary learning and teaching – in line with public intellectual ideals (Andrew, 2024) – can help cultivate a more knowledgeable, critical, and creative academic community. In doing so, we not only honour the legacy of literary greats like Joyce but also reaffirm our commitment to nurturing the full spectrum of human intellectual and creative potential.

Joyce is as inextricably associated with modern prose as Eliot is with modern poetry and Picasso with modern art (Ellman, 1982). In *Ulysses*, Joyce fundamentally reimagined the novel as a literary form. He pioneered a new modernist literature that articulated the perceived pointlessness and disorder that characterised early 20th-century Europe, and "dowdy, dirty Dublin" in particular (Birmingham, 2014, p. 54). Joyce's "revolutionary redefinition" of the novel in *Ulysses* treated all varieties of language – from the vernacular of Dublin to biblical references, from advertising slogans to classical allusions – as integral components of his expansive literary endeavour (Hastings, 2022, p. 72).

One must remember how stringent literary conventions were to fully appreciate Joyce's radical departure from them. A decade prior to the publication of *Ulysses*, Joyce faced seemingly insurmountable hurdles in publishing his short-story collection *Dubliners* (Joyce, 2008b) partly due to his use of the word "bloody", highlighting the extent of censorship and societal constraints on language (Birmingham, 2014, p. 225). Joyce's fearless incorporation of the F-word in *Ulysses* signalled a shift towards unfettered expression, eliminating previously inviolable taboos against freely articulating one's thoughts or ideas. Thus, Joyce's act of writing the word "fuck" in *Ulysses* transcended mere juvenile provocation.

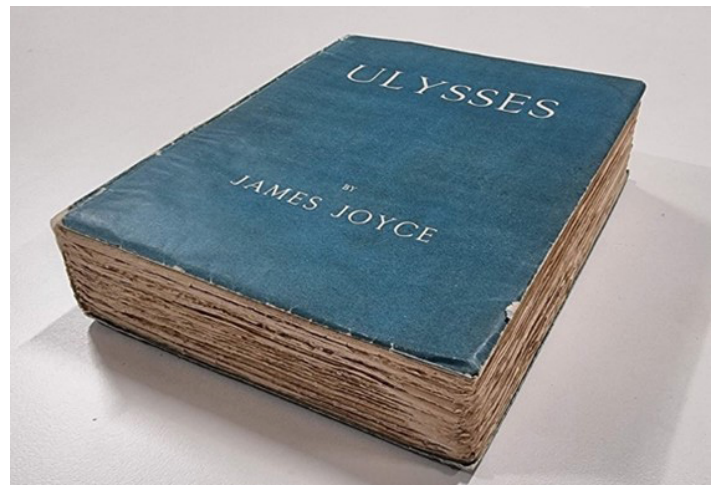


Figure 1. The first edition of *Ulysses* (1922). Bound in the Greek colours that Joyce considered lucky – white letters on a blue field – the book's design evokes the myth of Greece and Homer, reminiscent of a white island emerging from the sea. The formidable tome spans 732 pages, is three inches thick, and weighs nearly three and a half pounds (Ellman, 1982; Birmingham, 2014). Photograph of No. 302 of a limited edition of 1,000 numbered copies held by the State Library of New South Wales RB/0131.

Joyce's employment of profanities was merely a facet of a broader endeavour to dismantle established literary conventions. *Ulysses* challenged and deconstructed the conventional framework of narrative, offering a liberation from the established "tyranny of style" (Birmingham, 2014, p. 225). It marked a radical departure from established narrative techniques, dispensing with traditional narrative devices such as a singular narrative voice and blurring the lines between internal thought and the external world (Birmingham, 2014). In the "Oxen of the sun" chapter, Joyce crafts 32 parodies that trace the evolution of literary style from ancient pagan chants through Middle English, followed by the Latinate styles of Milton, imitations of satirists like Swift, and 19th-century novelists like Dickens. This chapter showcases Joyce's virtuosity in various writing styles, simultaneously advancing his narrative and paying ironic homage to the literary traditions that culminated in *Ulysses* (Hastings, 2022).

Already during its serialisation, *Ulysses* encountered fierce resistance from governmental bodies and various moral guardians, all zealous in their efforts to expurgate any perceived improprieties from literature. The novel's unabashed use of language and its bold, often provocative wordplay elicited considerable outrage. In addition, its irreverent portrayal of the British royal family and its 'blasphemous' views concerning the Roman Catholic Church intensified the scandal. Consequently, *Ulysses* found itself at the centre of legal battles and faced widespread censorship across the English-speaking world during the interwar period.

Despite these challenges, *Ulysses* has ascended to become one of the most highly regarded novels of the 20th century. 16 June 1904, the day Joyce and his future wife, Nora Barnacle, had their first date and the day the book's action takes place,



is commemorated globally every year as Bloomsday through festivities in bookshops and pubs (Hastings, 2022). More than 300 books and more than 3,000 scholarly articles are devoted, partly or entirely, to *Ulysses* (Birmingham, 2014). *Ulysses* continues to sell 100,000 copies a year, and it has been translated into more than 20 languages, including two Chinese translations (Birmingham, 2014). Random House continues to publish two rival editions of *Ulysses*.

One astonishing and innovative feature of *Ulysses* is its focus on a single day, evoking comparisons with the TV series "24", where each season's 24 episodes encapsulate an hour of a day in 'real-time'. McNamara (2010) described the series as akin to an "epic poem", with counter-terrorist federal agent Jack Bauer, a modern-day Odysseus, battling against human monsters that represent political corruption, cowardice, narcissism, megalomania, and terrorism. Contrary to "24", not much happens in *Ulysses*. Here is a feeble attempt to describe the content of the book in a paragraph.

Stephen Dedalus's day is filled with diverse activities: he shares breakfast with his roommates, teaches a class, enjoys a leisurely walk, engages in deep discussions with fellow intellectuals, indulges in alcohol, visits a brothel, and ends up being assaulted by an aggressive British soldier, a reflection of Ireland's status under British rule. Stephen is burdened by the memory of his mother's recent passing, a grief compounded by his refusal to pray for her due to his disenchantment with Catholicism. Meanwhile, Leopold Bloom starts his day by preparing breakfast for himself and his wife, Molly, before carrying out various errands around town. His day includes attending a funeral, conducting business in his role in advertising, dining out, engaging in a heated political debate with an Irish nationalist at a pub, spending time on the beach at dusk, and visiting a maternity hospital to check on a friend in labour. His path intersects with a drunken Stephen, prompting Bloom to take him under his wing. Bloom's day is complicated by his knowledge of Molly's afternoon affair with another man.

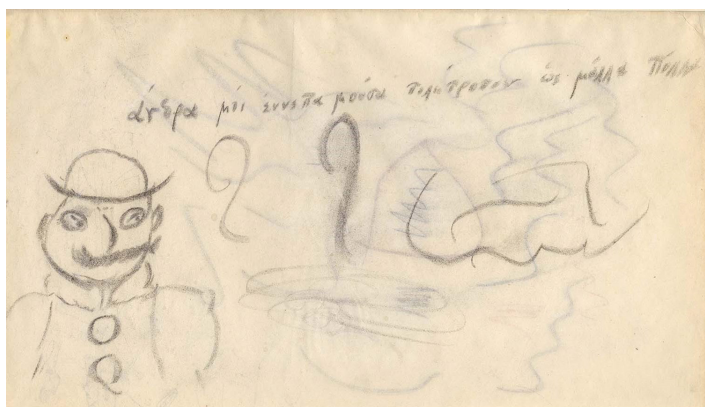


Figure 2. Joyce's sketch of Bloom is his only known visual depiction of one of his characters. It is a highlight of the McCormick Library's collection of 20th-century literary material. Source: <https://sites.northwestern.edu/northwesternlibrary/2018/06/05/collection-highlight-james-joyces-sketch-of-leopold-bloom/>

The initial section of this editorial continues with brief reconstructions of *Ulysses*'s critical reception, its parallels with Homer's *Odyssey*, and its unbelievably arduous journey through writing, printing, publishing, and censorship. Subsequently, we explore the stark disparity between such human-created masterpieces and the over-hyped generative artificial intelligence, including its algorithms' underwhelming tendency towards censorship. We then ponder the implications of these insights for effective teaching practices and leadership in higher education. Importantly, the second part of the editorial unveils and examines our latest issue, Volume 7(1).

### Critical reception of *Ulysses*

Joyce harboured a deep-seated belief that literature served as a testament to the resilience of the human spirit (Ellman, 1982). He maintained that as long as he could engage in writing, his physical surroundings were inconsequential, akin to Diogenes living in a tub (Ellman, 1982). Joyce had tremendous self-belief. He famously declared that he expected his readers to dedicate their entire lives to the study of his works (Eastman, 1931). Joyce infused *Ulysses* with myriad riddles and mysteries, rather accurately predicting it would occupy scholars for generations to debate his intentions, thereby securing his "immortality" (cited in Gifford, 1988, p. v). Joyce boldly proclaimed that "if *Ulysses* isn't fit to read, life isn't fit to live" (cited in Hutchins, 2016, p. 139).

To say that not everybody agreed with Joyce would be a gross understatement. Even D. H. Lawrence – who, in 1928, published his own 'obscene' novel *Lady Chatterley's lover* (Lawrence, 2006) that was banned till 1959 in the U.S. – described the final Penelope episode in *Ulysses* as "the dirtiest, most indecent, obscene thing ever written" (cited in Potter, 2009, p. 92). A book review in *The Daily Express* denounced *Ulysses* as "the maddest, muddiest, most loathsome book issued in our own, or any other time – inartistic, incoherent, unquotably nasty – a book that one would have thought could only emanate from a criminal lunatic asylum" (Mais, 1922). Another scathing review described *Ulysses* as

the most infamously obscene book in ancient or modern literature... All the secret sewers of vice are canalized in its flood of unimaginable thoughts, images and pornographic words. And its unclean lunacies are larded with appalling and revolting blasphemies directed against the Christian religion and against the name of Christ – blasphemies hitherto associated with the most degraded orgies of Satanism and the Black Mass (Douglas, 1922, p. 5).

*The Dublin Review* disapproved of the novel even more, condemning it as a "devilish drench" and calling upon the government to destroy the book (cited in Deming, 2013, p. 201). It also appealed to the Vatican to include it in the Index Expurgatorius, arguing that merely reading *Ulysses* was tantamount to sinning against the Holy Ghost – the sole sin deemed unforgivable by God's mercy (Birmingham, 2014). The famous psychologist C. G. Jung initially suspected

Joyce of being schizophrenic (Ellman, 1982) before revising his views and exclaiming that *Ulysses* was an alchemical laboratory that distilled “a new, universal consciousness” (Jung, 1979, p. 132).

Like Jung, literary luminaries such as Virginia Woolf and William Butler Yeats revised their initial negative assessments over time. Woolf (1923) initially described *Ulysses* as “a memorable catastrophe—immense in daring, terrific in disaster”. She further criticised the “illiterate, underbred book” as the effort of a “self-taught working man”, embodying the distressing characteristics of being “egotistic, inconsistent, raw, striking & ultimately nauseating”, going as far as to liken Joyce to “a queasy undergraduate scratching his pimples” (Woolf, 1980, pp. 188-189). However, by 1924, Woolf (2018) herself had authored *Mrs. Dalloway*, a novel obviously indebted to *Ulysses* as it explores the inner lives of its characters over a single day in London. Yeats initially dismissed *Ulysses* as “a mad book”, only to later concede, “I have made a terrible mistake. It is a work perhaps of genius... It is an entirely new thing... he has certainly surpassed in intensity any novelist of our time” (cited in Ellman, 1982, pp. 529-530).

Many literary greats unequivocally admired *Ulysses*. T. S. Eliot, who published *The wasteland* in the same year as Joyce *Ulysses*, wrote: “I hold this book to be the most important expression which the present age has found; it is a book to which we are all indebted, and from which none of us can escape” (Eliot, 1975, p. 175). F. Scott Fitzgerald, the author of *The great Gatsby* (originally published in 1925), offered to jump out a window to prove his devotion to Joyce and *Ulysses* – the offer, thankfully, was declined (Birmingham, 2014). Novelist Vladimir Nabokov (1990, p. 55) called *Ulysses* a “divine work of art” and the greatest masterpiece of 20th-century prose. Henry Miller compared the end of *Ulysses* to the end of the Book of Revelation (Birmingham, 2014). Hemingway swore that “Joyce has a most goddamn wonderful book” (cited in Ellman, 1982, p. 529) and wrote that “Jim Joyce was the only alive writer that I ever respected... he could write better than anyone I knew” (cited in Birmingham, 2014, p. 234).

### ***Ulysses* and *The Odyssey***

Thanks to Joyce’s student Borach, we have his teacher’s thoughts on *The Odyssey*.

“The most beautiful, all-embracing theme is that of the *Odyssey*.” It is greater, more human, than that of *Hamlet*, *Don Quixote*, *Dante*, *Faust*... I find the subject of *Ulysses* the most human in world literature. *Ulysses* didn’t want to go off to Troy; he knew that the official reason for the war, the dissemination of the culture of Hellas, was only a pretext for the Greek merchants, who were seeking new markets... I am almost afraid to treat such a theme; it’s overwhelming (Borach, 1917, cited in Ellman, 1982, pp. 416-417).

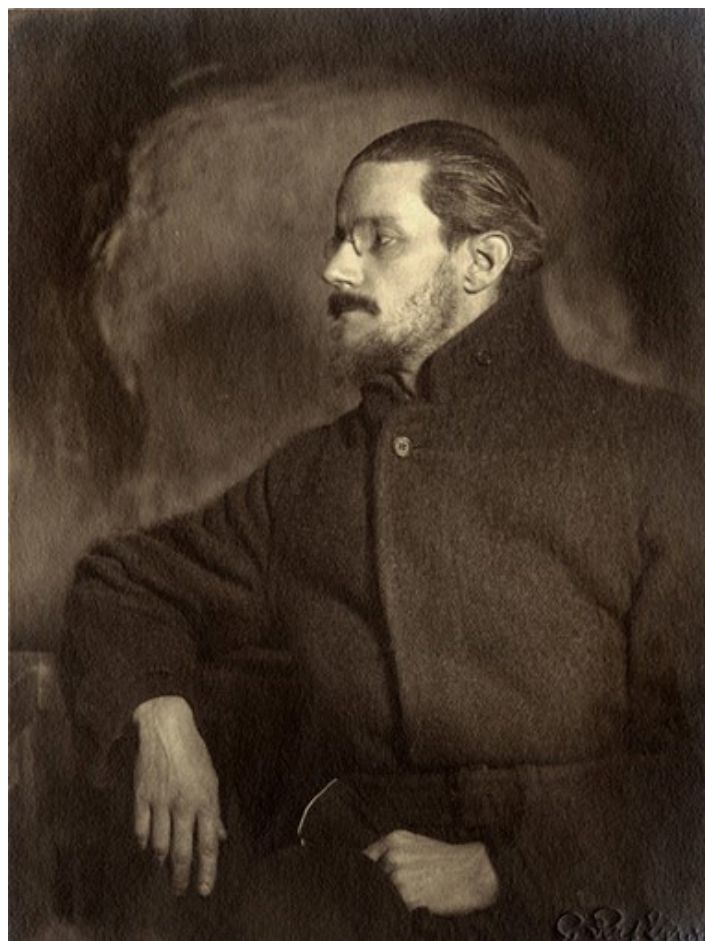


Figure 3. Photograph of Joyce by Camille Ruf, Zurich, ca. 1918. Cornell Joyce Collection, public domain.



Figure 4. *Odysseus and the Sirens*. Detail from an Attic red-figure stamnos by the Siren Painter (eponymous vase), circa 480-470 BC. Origin: Vulci. Public domain.

The 18 chapters of *Ulysses* roughly correspond to the 24 episodes in Homer’s *Odyssey* but are not in the original order. In Homer’s epic, Odysseus, a hero of the Trojan War, spends ten years journeying from Troy back to his home in Ithaca, facing tempests, a shipwreck, giants, monsters, and deities. To Joyce, Odysseus was simultaneously Europe’s first gentleman, a rebel and an ingenious warrior who came up with the first tank – the Trojan horse (Birmingham,

2014). Conversely, Joyce's novel unfolds over a mundane day in early 20th-century Dublin. Leopold Bloom, a Jewish advertising solicitor, corresponds to Odysseus. Stephen Dedalus, central to Joyce's prior semi-autobiographical work, *A portrait of the artist as a young man* (Joyce, 2000b), parallels Telemachus, Odysseus's son, while Bloom's wife, Molly, represents Penelope, who awaited Odysseus's return for two decades. Joyce playfully subverts these classical parallels. Ulysses is not a king but a newspaper advertising solicitor, and his homecoming is not to a loyal queen. While Penelope is renowned for her loyalty in warding off her suitors, Molly Bloom betrays her husband by engaging in an affair. Joyce invoked classical comparisons to critique the state of Western civilisation, represented through a single June day in drab, dingy Dublin.

To illustrate Joyce's ironic references to Homer, let's examine the "Cyclops" chapter (12). In Homer's (2006) epic (Book 9: "In the one-eyed giant's cave"), Odysseus and his men arrive at an island and venture into a cave. They encounter the cyclops Polyphemus, who devours some of Odysseus's crew and imprisons the survivors. Odysseus introduces himself as "Nobody" (Homer, 2006, 9.410) to Polyphemus, intoxicating the Cyclops and blinding him with a burning stake. When Polyphemus seeks aid, claiming, "Nobody's killing me" (Homer, 2006, 9.455), his kin conclude that if 'nobody' is the assailant, he must be suffering from a plague and refuse to assist. As Odysseus makes his escape, his unfortunate pride leads him to reveal his true name, prompting the enraged Polyphemus to throw a boulder at him and beseech Poseidon, his father, to curse Odysseus's voyage:

'Hear me—  
 Poseidon, god of the sea-blue mane who rocks  
 the earth!  
 If I really am your son and you claim to be my  
 father—  
 come, grant that Odysseus, raider of cities,  
 Laertes' son who makes his home in Ithaca,  
 never reaches home. Or if he's fated to see  
 his people once again and reach his well-built  
 house  
 and his own native country, let him come home  
 late  
 and come a broken man—all shipmates lost.  
 alone in a stranger's ship—  
 and let him find a world of pain at home!  
 (Homer, 2006, 9.585-595).

In Joyce's schema for *Ulysses*, the technique listed for this episode is "gigantism" (see Table 1). Accordingly, the Irish nationalist called "the Citizen" (Homer's Cyclops) becomes a "broadshouldered deepchested stronglimbed frankeyed redhaired freelyfreckled shaggybearded widemouthed largenosed longheaded deepvoiced barekneed brawnyhanded hairylegged ruddyfaced sinewyarmed hero" with "rocklike mountainous knees" whose "heart thundered rumblingly causing the ground, the summit of the lofty tower and the still loftier walls of the cave to vibrate and tremble" (Joyce, 2008a, 12.151-67). Odysseus's taking on the Cyclops with a burning stake in the cave is hilariously paralleled by Ulysses arguing with the bigoted Citizen in a pub while brandishing a lit cigar.

Eventually, the Citizen, brimming with racist animosity towards Bloom, initiates a confrontation. As Bloom departs from the pub, the Citizen follows and taunts him with a derisive cheer for Israel. Bloom counters this by citing renowned Jewish philosophers and artists, including Jesus. This provocation maddens the Citizen, who storms back into the pub, seizes a biscuit tin, and hurls it at Bloom as his carriage pulls away. The scene is depicted with the intensity of a seismic event and portrayed with "imagery of biblical rapture" (Hastings, 2022, p. 145).

In view of the vast size and intricate complexity of *Ulysses*, Joyce eventually permitted the publication of his "summary-key-skeleton-scheme" (cited in Hastings, 2022, p. 277). The schema maps out each chapter with specific times (on June 16 and 17, 1904), settings, colour schemes, narrative techniques, parallels (mostly to *The Odyssey*), related fields of science or art, thematic significances, associated human organs, and symbols. The schema for the "Cyclops" chapter is presented in Table 1 as an example.

Table 1. Joyce's schema for the Cyclops episode. Adapted from Hastings, 2022, pp. 280-281.

Episode	Cyclops
Time	5 p.m.
Scene	Tavern
Colour	Green
Technique	Gigantism
Correspondences	I = Noman Cigar = Stake Apotheosis = Challenge
Science/art	Politics
Meaning	The egocidal terror
Organ	Muscles
Symbols	Nation, Religion, Gymnastics, Idealism, Exaggeration, Fanaticism, Collectivity

### Writing, printing, publishing, censoring, burning, and celebrating *Ulysses*

The heaventree of stars hung with humid  
 nightblue fruit  
 (Joyce, 2008a, 17.1039).

*Ulysses* had to navigate perilous paths to be written, printed, published and disseminated. When Joyce started to write *Ulysses* in 1914, he began his own Odyssey. At multiple points, the world came close to never seeing this now-classic masterpiece. The fact that Joyce managed to write, print, and publish *Ulysses* at all is miraculous. Even with his comparatively tame previous works, he had faced the greatest problems that would have made lesser authors give up. Already in the 1900s, Joyce realised that "I cannot write without offending people", and his career reminded him "of an opera with a magnificent overture... While the audience is applauding just before the curtain goes up, in comes a group of bumbailiffs and arrests the fiddlers for debt" (cited in Ellman, 1982, pp. 210, 264).

Even before *Ulysses*, printing his works had been extremely trying. Joyce's short-story collection *Dubliners* required his nine-year-long "correspondence with seven solicitors, one hundred and twenty newspapers, and several men of letters about it – all of whom, except Mr Ezra Pound, refused to aid me" (Joyce, cited in Ellman, 1982, p. 415). The first edition

(1906) was aborted, the second burnt (1910), and the third (1914) was eventually published after 40 publishers had rejected it.

His novel *A portrait of the artist as a young man* (Joyce, 2008b) was refused by every publisher. Joyce's sense of powerlessness fuelled his resentment. In a moment of despair, he cast the incomplete manuscript into the fire. Thankfully, Joyce's sister Eileen retrieved the manuscript from the flames at the cost of minor burns to her fingers (Ellman, 1982). Later, Joyce considered buying a revolver and putting "some daylight into my publisher" (cited in Ellman, 1982, p. 331). Instead, he wrote the amusing broadside "Gas from a burner", ostensibly spoken by his publisher and printer:

Ladies and gents, you are here assembled  
To hear why earth and heaven trembled  
Because of the black and sinister arts  
Of an Irish writer in foreign parts  
He sent me a book ten years ago;  
I read it a hundred times or so,  
Backwards and forwards, down and up,  
Through both ends of a telescope.  
I printed it all to the very last word  
But by the mercy of the Lord  
The darkness of my mind was rent  
And I saw the writer's foul intent.  
But I owe a duty to Ireland:  
I hold her honour in my hand,  
This lovely land that always sent  
Her writers and artists to banishment  
(cited in Ellman, 1982, p. 336)

When *The Egoist* finally decided to publish *A portrait of the artist*, about twenty printers in England and Scotland refused to print it (Ellman, 1982).

Joyce had laid the groundwork for *Ulysses* since 1907, with the project becoming increasingly ambitious in scope and method over time (Ellman, 1982). He estimated that he had dedicated around 20,000 hours (Ellman, 1982) to writing approximately 265,000 words for *Ulysses*. That is a glacially slow average 'writing' speed of 13-14 words per hour (that presumably includes researching and editing). For instance, he devoted an entire day to perfecting these two sentences (Hastings, 2022): "Perfume of embraces all him assailed. With hungered flesh obscurely, he mutely craved to adore" (Joyce, 2008a, 8.638-39). Joyce described a state of total mental exhaustion following the completion of an episode, feeling as though neither he nor the "wretched book" would recover from the effort (cited in Ellman, 1982, p. 461). He worked "like a galley-slave, an ass, a brute", could not even sleep, and the "episode of *Circe* has changed me too into an animal" (Joyce, cited in Birmingham, 2014, p. 180).

Even Ezra Pound, Joyce's staunch supporter, wrote to ask if he had "got knocked on the head or bit by a wild dog and gone dotty" (cited in Birmingham, 2014, p. 132) when reading the beginning of the *Sirens* episode. We cite a couple of lines for your enjoyment.

Bronze by gold heard the hoofirons, steelyringing,  
Imperthnthn thnthnthn...  
Jingle jingle jaunted jingling.  
Coin rang. Clock clacked.  
(Joyce, 2008a, 11.1-2, 15-16)

Here are some hints: Bronze and gold were the principal metals in Homer's epic; Miss Douce, a barmaid, threatens to report a customer for his "impertinent insolence" that is parodied by the busboy as "Imperthnthn thnthnthn"; a "jingle" and a "jaunting car" are two-wheeled horse-drawn carriages; the clock strikes 4 p.m. (see Gifford, 1988, pp. 290-291, 86).

Joyce continued to work on *Ulysses* almost up to its publication day, with about a third of the novel being written during the proofreading stage (Ellman, 1982). Joyce requested five sets of proofs and, using his notes, made countless alterations – primarily expansions. He felt compelled to engage in an exhaustive routine of writing, revising, and correcting for approximately twelve hours daily, with brief pauses when his vision blurred (Ellman, 1982).

Chronic severe eye conditions significantly hampered Joyce's literary endeavours and daily life. He suffered from recurrent iritis (inflammation of the iris), leading to episodes of acute glaucoma and additional complications that drastically diminished his vision, nearly to blindness. Joyce's treatments were as harrowing as his symptoms. He endured not only the prospect of having his eyes "slit open" but also a relentless regimen of injections, narcotics, antiseptics, dental extractions, and the application of tonics, electrodes, and leeches (Birmingham, 2014, p. 9). Considering the agony Joyce experienced, it is astonishing that he managed to write *Ulysses*.

Moreover, the book's creation coincided with the tumultuous era of the Great War and its aftermath. Spanning from 1914 to 1918, the First World War claimed the lives of 17 million people, both military and civilian, and unveiled a "monstrous epiphany in the European imagination" (Birmingham, 2014, p. 59). The aftermath of this catastrophic period brought the Spanish flu pandemic, which killed as many as 100 million people worldwide (Barry, 2020).

In Joyce's view, artists should eschew judging their characters through a moral lens, instead approaching even the most deviant behaviours with "indifferent sympathy" (cited in Ellman, 1982, p. 139). Joyce viewed his work as a meticulously buffed mirror reflecting reality. If the reflection appears repugnant, the fault lies not with the mirror (Hastings, 2022). Joyce held the principle of free expression in high regard. To him, censorship represented an overreach of governmental authority, dictating not only the prohibition of obscenity but also defining what was 'obscene' in the first place (Birmingham, 2014). Joyce (cited in Ellman, 1982, p. 688) commented on the "strangely hostile reception" of *Ulysses* and his own perspective of morality:

The most natural thing for a writer is to call a spade a spade. The mistake that some moralists make, even today, is that they hate unpleasant phenomena less than they do those who record them. It's always the

same. People go on judging an author immoral who refuses to be silent about what in any case exists. Immoral! Why, it's a mark of morality not only to say what one thinks is true—but to create a work of art with the utmost sacrifice; that's moral, too.

*Ulysses* was serialised in the American journal *The Little Review* from 1918 to 1920. Birmingham recounts the amusing anecdote where British war censors believed the serialised parts of *Ulysses* to be a complex spy code (Birmingham, 2014). Birmingham (2014), in his brilliant *The most dangerous book: The battle for James Joyce's Ulysses*, shows that *Ulysses's* initial difficulties with censorship did not stem from vigilantes hunting for pornographic content but rather from Post Office government censors on the lookout for foreign spies, radicals, and anarchists.

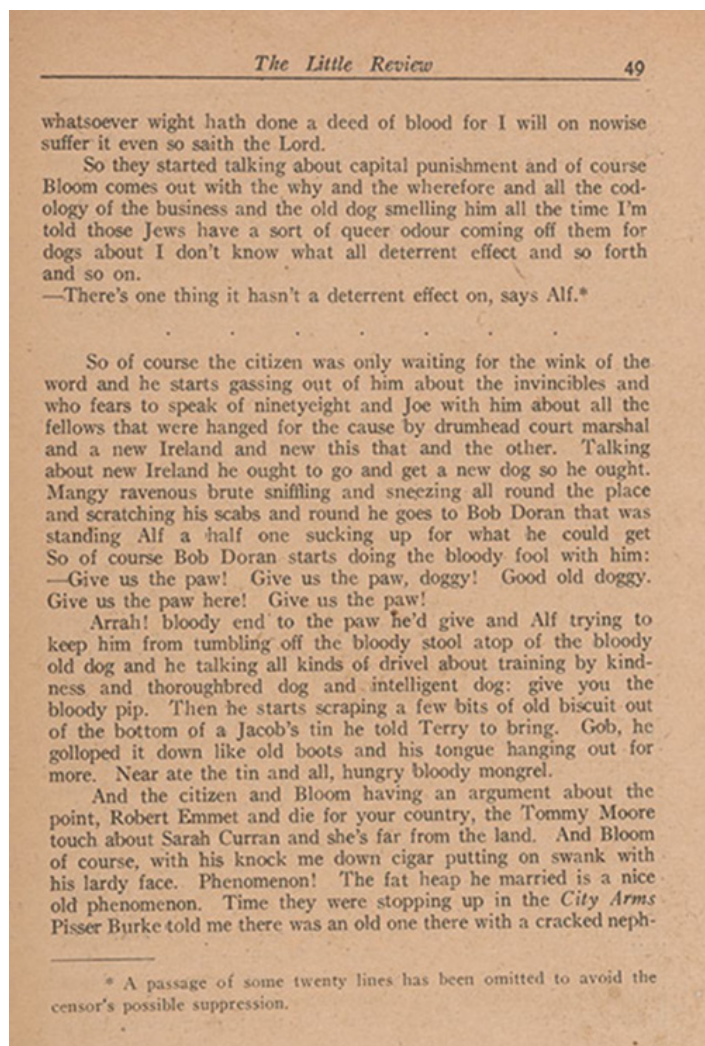


Figure 5. Self-censorship of parts of the "Cyclops" episode in *The Little Review*, 6(7), November 1919. "Cyclops", as Joyce originally wrote it, never appeared in the magazine in its entirety. Following the US Postal Service's suppression of two *Ulysses* issues, the editors pre-emptively censored the first instalment of "Cyclops.". Passages were replaced with an asterisk, an ellipsis, and a telling footnote. The Morgan Library & Museum, gift of Sean and Mary Kelly, 2018; PML 197868.8. Public domain.

The Comstock Act of 1873 criminalised the distribution of materials considered obscene material via the U.S. postal service. Legislation prohibiting obscene literature emerged in the mid-19th century, driven by increasing literacy and urban growth. This era's heightened concern for public morality led to the stringent application of such laws, and *Ulysses* was an ideal subject for enforcement (Birmingham, 2014). Consequently, seized editions of *The Little Review* that contained initial versions of chapters from *Ulysses* were handed over to the Salvation Army, where women in reform programmes were tasked with ripping them to shreds (Birmingham, 2014). In an eery anticipation of the Nazi book burnings, officials collected nearly 500 copies of *Ulysses* and disposed of them in the furnace room of a post office building. The Comstock Act enraged Ezra Pound, who was instrumental in getting *Ulysses* serialised in the US. He lambasted it as "grotesque, barbarous, ridiculous, risible, Gargantuan, idiotic... pissian, phartian, monstrous, aborted, contorted, distorted, merdicious, stinkiferous, pestilent" (cited in Birmingham, 2014, p. 119).

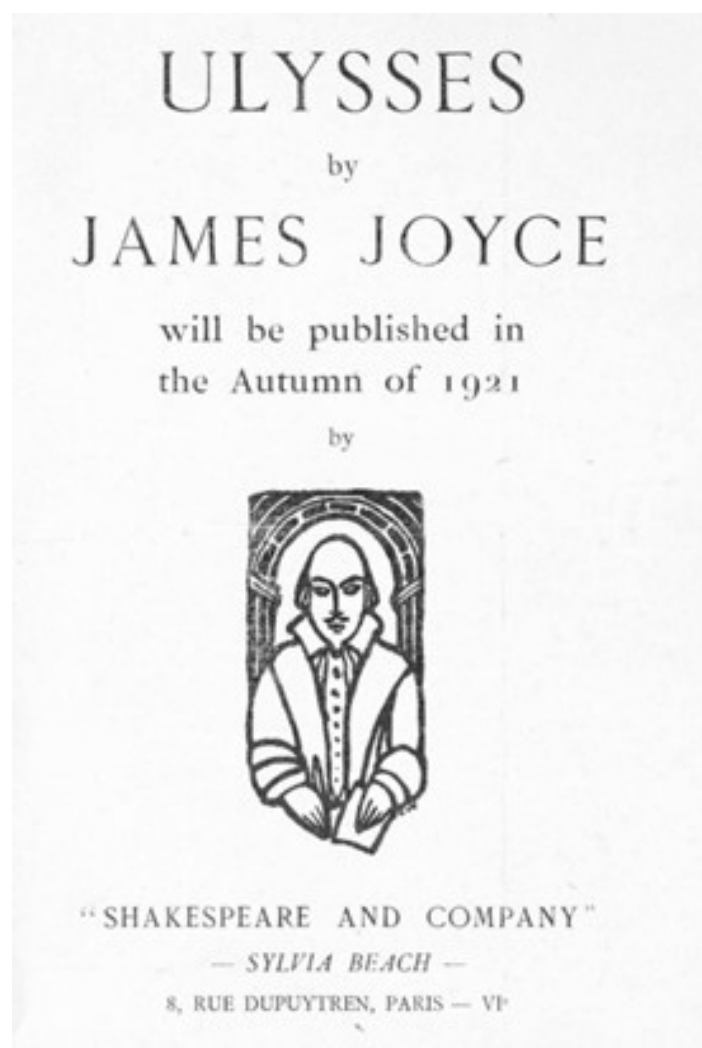


Figure 6. Poster announcing the publication of *Ulysses*. The first copies were only to become available on Joyce's 40th birthday, 2 February, 1922. Public domain.

Eventually, Judge Woolsey presided over the most publicised obscenity case in U.S. history in 1933, *United States v. One book called Ulysses*. The explicit content in *Ulysses* exceeded anything previously permitted by American courts. However, for every obscene term, there were numerous obscure ones,

like 'quadrireme', 'entelechy', 'epicene', or 'hebdomodary'. *Ulysses* was a masterpiece to the literati, while it appeared incomprehensible to those deemed 'morally at risk' (Birmingham, 2014).

Judge Woolsey recognised that for Joyce, all elements, including propriety, were secondary to his artistic vision. He ruled that Joyce, despite facing criticism and misunderstanding, "has been loyal to his technique" and "has honestly attempted to tell fully what his characters think about", regardless of the consequences. Woolsey did more than merely legalise Joyce's work; he championed it, stating, "Ulysses is an amazing tour de force" (cited in Birmingham, 2014, p. 329). *Time* lauded Woolsey's decision as "historic for its authority, its eloquence, its future influence on U.S. book publishing" (cited in Birmingham, 2014, p. 330).



Figure 7. Family portrait photograph of the Joyce family. Clockwise from top left: James Joyce, son Giorgio Joyce, wife Nora Barnacle, and daughter Lucia Joyce (1 January 1924, unknown photographer). Public domain.

In the 1920s, U.S. and British government agencies had incinerated numerous copies of a book that, by the 1930s, had emerged as a modern classic and an integral part of Western cultural heritage. The official acceptance of *Ulysses* signified that the experimental and radical culture of the 1910s and 1920s had not been a mere deviation but had firmly established itself in literature. By allowing *Ulysses* to be published, authorities acknowledged the fluidity of cultural norms, admitting that distinctions between what is considered 'classic' and what is deemed 'filth' were not rigid

(Birmingham, 2014). It is remarkable to think that, less than a hundred years ago, powerful forces in the United States were determined to stifle artistic expression in the guise of safeguarding public morals. Their efforts inadvertently shifted *Ulysses* from being the flagship of avant-garde literature to a broader symbol of artistic freedom. A world devoid of *Ulysses* would be significantly diminished.

### The inferiority of generative AI and consequences for good teaching and higher education leadership

Amidst the surrounding hyperbole, it is clear that the current generation of Large Language Models (LLMs), including GPT-4o and its rivals, falls markedly short of capturing the profound depth and intricate richness embodied by literary classics like Joyce's *Ulysses* or Homer's *Odyssey*. It is worthwhile recalling that *The Odyssey* was written 2,700 years ago. For those who have truly immersed themselves in Homer's exceptional *Iliad* and *Odyssey*, the notion of generative AI producing work that even remotely approaches the magnificence of these ancient epics is fanciful. Such a notion underscores a profound superficiality that derives from insufficient engagement with literature and even non-fiction books (Sam, 2024). It conflates the oftentimes terminally dull prose of GenAI with the literary classics that we have exemplified through *Ulysses* and *The Odyssey*. Techno-optimistic and solutionist claims for generative AI are repeated ad nauseam, and there is a dearth of critical voices (Lindgren, 2023a, 2023b; Popenici et al., 2023; Rudolph et al., 2024b). One does not need to live in Nazi Germany or in contemporary Russia to know that when lies are sufficiently frequently repeated, they become truths (Rudolph & Tan, 2022). Or, to cite the wonderful title of Pomerantsev's (2017) book: *Nothing is true and everything is possible*.

Human intelligence is characterised by multiplicity and far superior to the ill-named 'artificial intelligence'. Academics babbling endlessly about "Artificial General Intelligence" (AGI) or "superintelligence" (e.g. Bostrom, 2017, Kurzweil, 2005) – not to speak of Elon Musk predicting that "AI will overtake human intelligence next year" (Hammond, 2024) – usually have a vested interest in these claims. Tech tycoons and their allied 'thought leaders' benefit from singing from the techno-optimistic and solutionist hymn sheet. However, the expectation that 'newer equals better' in every aspect oftentimes falters, not only in the realm of creative literary expression.

These days, fewer people appreciate physical books (Sam, 2024), and we wish you the best of luck finding anybody with the verbal prowess of Joyce or Homer among your acquaintances. There is popular disdain for bibliophiles, and to you, book lovers and collectors, we give you Walter Benjamin's quote (first published in 1931), showing that this contempt is nothing radically new:

Suffice it to quote the answer that Anatole France gave to the philistine who admired his library and then finished with the standard question, 'And you have read all these books, Monsieur France?' 'Not one-tenth of them. I don't suppose you use your Sèvres china every day?' (Benjamin, 2015, p. 64).

There is a clear and present danger that we are becoming dumber and dumber while machines are getting smarter and smarter (Popenici, 2023). As educators, it is our job to expose that 'artificial intelligence' is neither artificial nor intelligent (Crawford, 2021), combat shallowness (Carr, 2020) that mistakes quantity for quality and that succumbs to "garbage in, garbage out" outputs in a "new dark age" (Bridle, 2023; Rudolph et al., 2024b). Many experts warn against equating human and artificial intelligence, as human intelligence cannot be fully captured in precise, machine-compatible descriptions (Verdicchio, 2023; Luckin et al., 2024). Gardner (1993) influentially conceptualised multiple intelligences, which categorises human intelligence into eight distinct types: visual-spatial, linguistic-verbal, logical-mathematical, bodily-kinaesthetic, musical, interpersonal, intrapersonal, and naturalistic. If Gardner's multiple intelligences are agreed upon, and intelligence is not reduced to its logical-mathematical aspect, then AGI would have to encapsulate all eight intelligences.

In terms of censorship, the lifting of the publishing bans on *Ulysses* in the 1930s was spectacular and thrilling to liberals. Alas, such hard-fought freedoms do not last automatically, and AI plays a sinister role in strengthening censorship. Authoritarian regimes have adapted to AI chatbot innovations by aligning them with their censorship agendas, utilising machine learning on digital platforms to filter out unwanted political, social, and religious discourse. The adoption of AI has amplified the scope and effectiveness of digital suppression, enabling online censorship through automated systems. Advanced surveillance technologies swiftly scour social media for dissent, combining large data pools with facial recognition to monitor and identify pro-democracy activists (Freedom House, 2023). The World Economic Forum (2024) has labelled AI-generated disinformation the most severe global threat at present.

While there is consensus on preventing GenAI from facilitating the creation of biological or chemical weapons or child pornography, leading GenAI chatbots, such as ChatGPT and Gemini, implement broad and ambiguous 'harm' filters "that leave users in the dark about where, how, and why the red lines are drawn" (Mchangama & White, 2024). For example, GenAI's refusal to detail comedian Lenny Bruce's (1925 - 1966) controversial yet impactful work illustrates the excessive caution in content moderation. ChatGPT maintained that it is unable to offer instances of "slurs, blasphemous language, sexual language, or profanity", committing instead to "share information in a way that's respectful and appropriate for all users". Gemini took an even more cautious stance, suggesting that presenting Bruce's language "without careful framing could be hurtful or even harmful to certain audiences" (Mchangama & White, 2024).

When we asked GPT-4 to provide reasons and examples as to why *Ulysses* was banned in much of the English-speaking world in the 1920s and early 1930s, it did well in elaborating on the four aspects of "Sexual Content and Obscenity", "Language and Profanity", "Challenging Conventional Morality and Social Norms", and "Anti-Authority and Anti-Religious Sentiments". When asked to cite some of the offensive passages, it explained that "I can't directly quote passages from '*Ulysses*' or any copyrighted material", but

volunteered to "guide you on where to find some of the controversial content", proceeding to elaborate on the Nausicaa, Circe and Penelope episodes in particular, and commenting on "Language and Profanity" as follows: "Joyce employs a wide range of language, including what was considered profanity and coarse language at the time. While specific instances are spread throughout the text, the overall use of such language contributed to the book's initial reception as obscene". Unsurprisingly, and in line with Mchangama and White's (2024) findings, ChatGPT does not cite any profanities.

The censorship described by Mchangama and White (2024) raises concerns about the extent of information and viewpoints being filtered out under the guise of harm prevention. They argue that GenAI must not replace human judgment and criticise the current approach by a small group of powerful companies for limiting open inquiry and expression based on vague and unsubstantiated claims of 'harm'. They caution against a future dominated by AI's restrictive moral frameworks in everyday technologies, advocating for access to a broad spectrum of information.

Mchangama and White (2024) paint the dystopian scenario where "your word processor prevents you from analyzing, criticizing, lauding, or reporting on a topic deemed 'harmful' by an AI programmed to only process ideas that are 'respectful and appropriate for all'". With the rapid integration of GenAI into search, word processing and email, this frightening prospect may not be as far-fetched as it initially sounds. Guardrails should avoid restricting human agency or curiosity. We need to think for ourselves and make more informed decisions based on a wealth of information from multiple perspectives. We need to ensure that AI systems are optimised to enhance human reasoning, not to replace human faculties with the "artificial morality" of large tech companies (Mchangama & White, 2024).

Amidst the burgeoning discourse around GenAI, it may be good to go back to basics and remind ourselves that good teaching means "being willing to do anything that helps students learn" (Brookfield et al., 2023, p. viii). Teaching well encompasses adapting our teaching strategies to align with the unique contexts we encounter. The selection of teaching modalities should be informed by our understanding of the context and our educational goals, choosing those that we believe will most benefit our students' learning journey. Enhancing our teaching practices requires us to embrace experimentation and calculated risk-taking, underpinned by continuous feedback from students (and ideally, peers) through persistent classroom research. Thus, teaching evolves "as a continuous process of failing well, in which our growing appreciation of complexity is matched by a willingness to be more and more open to different approaches" (Brookfield et al., 2023, p. ix). The teaching for a critical AI literacy and an emphasis on critical thinking, creativity and teamwork will be key (Rudolph et al., 2024a).

In our turbulent age, learning leadership in higher education will have to come from below, behind, and among, thus including multiple stakeholders such as teachers and students (Preskill & Brookfield, 2009; Rudolph et al., 2024a). It is essential to guide students on the ethical and critical

use of generative AI in their learning to enhance their critical reflection skills (Tan et al., 2024). As educators, we need to foster deeper relationships with our students, encouraging discourse while educating them to critique authoritative-sounding misinformation (Mills, 2023). Debunking anthropomorphic tendencies is essential in nurturing well-rounded, critical thinkers (Mills, 2023; Rudolph et al., 2023b). We advise higher education institutions to foster multi-stakeholder dialogues, including representatives from various sectors, to integrate the insights gleaned into concrete guidelines, regulations, and educational materials, emphasising the pivotal role of digital literacy education, which encompasses a range of AI tools (Gimpel et al., 2023; Rudolph et al., 2023a). There is a fast-growing literature on AI and higher education teaching and learning that Ismail et al. (2023, 2024a) have organised into a freely available open-access database (Ismail et al., 2024b) to facilitate critical discussions.

Finally, it is deeply ironic that much-maligned humanities may produce superior prompt engineers, commanding six-digit starting salaries (Marr, 2023). A combination of communication skills, critical thinking and creativity continue to be extremely useful. Joyce, if transported with a time machine into our age, would have likely avoided GenAI and despised its inherent censorship. Otherwise, he might have had a good laugh about the sudden market value of people who could write well.

## Overview of issue 7(1)

This issue, one of our most voluminous to date, once again shines a spotlight on artificial intelligence in higher education. It offers rich insights and practical perspectives on how various institutions navigate the complex AI landscape in academia. Leading the discussion is Professor Waring's thought-provoking commentary, "Artificial intelligence and graduate employability: What should we teach generation AI?" In his contribution, he explores the intersection of AI and graduate employability, advocating for curriculum adjustments that incorporate AI literacy and ethical decision-making. Waring concludes by championing a balanced approach, combining technical skills with critical thinking and interpersonal abilities, to equip graduates for an AI-driven job market.

Following Waring's commentary are 29 research articles that explore a diverse range of topics, from the ethical implications of GenAI in higher education to the development of virtual reality (VR) and digital literacy competencies. These articles also address pressing issues such as the challenges posed by neoliberalism, toxic supervision in PhD programmes, cultural safety education, and organisational resilience. In addition, a practical piece provides valuable guidance on leadership within higher education institutions.

The research section opens with a thematic exploration of education research after the COVID-19 pandemic by Bala and Mitchell. Their study employed BERTopic modelling to analyse trends and research within the *Journal of Applied Learning and Teaching*, revealing thematic structures and emerging trends. It identified 17 topics across four thematic

groups, reflecting global trends in post-COVID learning and teaching, and providing insights for future research and practice in higher education.

11 manuscripts in the section explored AI topics and higher education. First is Van Wyk's article, "Is ChatGPT an opportunity or a threat?" His study employed semi-structured interviews to investigate academics' perceptions of ChatGPT at an education faculty and the findings study highlights its potential for enhancing teaching and learning in preventing academic dishonesty.

Second, Ogunleye et al.'s "Higher education assessment practice in the era of GenAI tools" examined how GenAI tools affect higher education (HE), particularly assessment and academic practice. The study looked at three disciplines in particular (data science, data analytics, and construction management) and highlighted GenAI's potential benefits and limitations. The authors also advised the ethical use of GenAI and offered recommendations for integrating AI tools into higher education's teaching and learning.

This is followed by Ahmad et al.'s paper on AI tools among Asian and African higher education staff and students, concluding that awareness, benefits, threats, attitudes, and satisfaction are critical factors determining its usefulness. Among 815 participants, 38% were unaware of the presence and functionality of AI tools, but 63% revealed that they rarely use AI tools. Notably, higher education-level individuals perceive AI tools as being threatening, while female participants expressed more concerns than males. These findings underscore the diverse levels of comfort and familiarity with AI technology across demographics and educational backgrounds, emphasising the necessity of enhancing AI awareness and development in Asia and Africa.

Next comes Kouam and Muchowe's insightful piece on graduate students' perception and adoption of AI chatbots in Zimbabwe. They investigated Zimbabwean graduate students' perceptions and adoption of AI chatbots in universities and examined benefits like enhanced learning and skill development alongside challenges such as plagiarism and financial constraints. Findings revealed graduate students' positive attitudes towards chatbots, highlighting their role in augmenting education while recommending measures for better integration and accessibility.

Another paper by Joseph et al. similarly presented insights from a multi-group analysis of students' awareness and perceptions using gender and programme type concerning the use of AI tools for research. Male and postgraduate students demonstrated higher awareness and perception, while female students excelled in using AI tools for research. This study underscores the importance of incorporating AI tools into university curricula while considering demographic variables for technology integration.

This is followed by Sobaih's "Ethical concerns for using artificial intelligence chatbots in research and publication: Evidences from Saudi Arabia". This study surveyed academics and research leaders, revealing widespread chatbot use, ethical dilemmas, and pseudoscience risks. Strategies to



mitigate concerns and promote responsible usage were recommended, and the findings emphasised the critical need for effective policy interventions.

Kershnee and Potter's article on GenAI in distance education shed light on the challenges, and impact on academic integrity and student voice in distance education. Using the technology acceptance model, it investigates GenAI's influence on learning, integrity, and student voices in a South African open distance and e-learning university. Qualitative data from interviews, evaluations, and focus groups highlight the need to balance concerns with the potential benefits of GenAI in education.

Next, a systematic review by Chaka examined studies discussing the performances of different AI detection tools in differentiating between AI-generated and human-written text. The review screened 17 articles and concluded that the detection tools had varied detection efficacy, and suggested adopting a combined approach involving AI detectors, traditional tools, and human reviewers. Another related paper by Chaka, "Accuracy pecking order – How 30 AI detectors stack up in detecting generative artificial intelligence content in university English L1 and English L2 student essays", evaluated the accuracy of 30 AI detectors' accuracy in identifying GenAI-generated and human-written content in university English L1 and L2 student essays. Results showed that only two detectors, Copyleaks and Undetectable AI, accurately identified all essays as human-written. Most detectors misidentified the essays and were deemed unfit for purpose.

Lastly, Ismail et al. introduced an open-access repository using a systematic literature review: "Artificial intelligence in higher education database (AIHE V1)". Utilising a rigorous systematic review method, the review provides a first look at the metadata of articles published on AI and higher education during ChatGPT's inaugural year, facilitating scholars and practitioners in making informed decisions in relation to policy and practice. The open access database is freely accessible via a separate DOI (Ismail et al., 2024b).

The next three articles are related to VR simulation and digital competencies in higher education. The first is Abusalim et al.'s "Digital versus classroom discussions: Motivation and self-efficacy outcomes in speaking courses via Gather.town". The authors debate online and traditional classroom methods, focusing on student motivation and self-confidence in a German language-speaking course. Results demonstrated significant improvements in motivation and self-efficacy in the online group, highlighting the potential of platforms like Gather.town in enhancing educational outcomes.

Second, Inkabi et al.'s "Utilizing head simulation training in dental school education: Time and cost implications" investigates cost and time factors as barriers to the effectiveness of head simulator use in dental schools. Findings indicate that most participants disagreed that head simulators extend course duration. While their availability was generally rated positively, the cost of using these simulators did not significantly impact device accessibility or course duration.

Next, the article by Rojas-Osorio et al., "Self-perception of university teachers on their digital teaching competence: The case of Peru", analysed the self-perception of digital competence of 122 university professors at a private university in Peru. The study revealed low participation in training, evaluation, and innovation projects with ICT, indicating a need for continuous training programmes to enhance digital competence among professors.

Neoliberalism took centre stage in the next three articles. Martin Andrew's article "'Just get them over the line': Neoliberalism and the execution of 'excellence'" is grounded in the author's experience of postgraduate education using narrative inquiry to examine the concept of excellence in postgraduate education. He suggests that the neoliberal conception of 'excellence' hides a more authentic form of 'excellence' and believes that this form can only surface if the voices of learners and educators are heard above the managerialist chatter and when teaching well is considered. Martin Andrew's piece on 'The Great Resignation: The simple joys of not belonging' continued his argument of the harm the neoliberal grip has on higher education. Similarly, using vignettes as a form of narrative inquiry, his article explored the relational link between the archaic notion of affliction and what it means to 'belong' to a university for academics. The narratives revealed the importance of exercising critical resilience to establish academic identities beyond the neoliberal university. Nikpouya and Zareian's "Neoliberalism and the violation of students' rights: The case of English language education" completes the argument on the dangers and challenges of neoliberalism on higher education in this issue. Their theoretical study explores the impact of neoliberalism on education, focusing on areas of general education and the English language. They found that neoliberal ideologies have led to the commodification of education and amplified the emphasis on standardised testing and accountability measures.

The following two articles explore power relationships in academia. Owan et al.'s "Metrics in research impact assessment and grant funding: Insights from researchers in the "Reviewer 2 Must Be Stopped!" Facebook group" explores the reliance on quantitative metrics in research assessment and grant funding, gathering insights from over 15 experienced researchers worldwide. Data were analysed thematically, revealing diverse perspectives. While some voiced concerns about metric dominance and biases, others recognised their value. The study emphasises the need for a balanced, context-aware approach incorporating qualitative measures. Okere's "A content analysis of tweets on toxic doctoral supervision" investigates toxic supervision of PhD students via Twitter posts. A content analysis of these posts reveals themes and trends, shedding light on students' experiences. Twitter (recently renamed X) has emerged as a valuable research tool and support platform for doctoral researchers. The findings sought to inform policy and enhance supervisory practices in academia.

The subsequent research articles in this issue encompass a variety of topics. We start with Moore et al.'s "The challenge of making relationships central in online cultural safety education". It explores cultural safety education, emphasising the importance of fostering respectful classroom

relationships. Collaborative reflections by university educators compare facilitating positive connections in online versus physical classrooms. Findings reveal how technological affordances affect relational dynamics, impacting educators' emotional labour. The study suggests integrating culturally responsive pedagogies to prioritise relationship-building and support effective teaching across physical and online learning environments.

Williams's "A conceptual, strategic and implementation framework for the Scholarship of Learning and Teaching" explored the Scholarship of Learning and Teaching (SoLT) framework in an Australian higher education provider. It outlines standards for scholarly practice and explores the linkages between scholarship, research, professional development, and quality improvement. The initiative has shown success, with a high percentage of academics meeting standards for scholarly practice.

Hanshaw's article "Micro-credentials in higher and vocational education: An innovation or a disruption?" critically assesses the literature on micro-credentials in higher and vocational education, debating their potential as innovative tools or disruptive elements. It explores their role in promoting agency, equity, access, and their perceived simplification of academic credentials. Hanshaw advocates for the leveraging of micro-credentials to drive positive innovation in education.

Shafi and Middleton's "Organisational resilience in a higher education institution: Maintaining academic continuity, academic rigour and student experience in the face of major disruption (Covid-19 pandemic)" investigates a university's response to the Covid-19 pandemic in England. Data from various levels were collected using a systems-based approach and an organisational resilience framework. The study highlights the university's positive adjustments during the pandemic. Still, it underscores the need to understand longer-term impacts and resilience in adversity.

Calonge et al.'s "Do graduate courses in a HyFlex mode foster emotional, cognitive and behavioral engagement? A consideration" explored strategies and their impacts on learning outcomes and engagement. Through critical reflection, results showed that HyFlex courses can achieve equivalent learning outcomes but require staff development and purposeful activity design to promote emotional, cognitive, and behavioural engagement.

In the same vein, Khatter et al.'s "Student engagement and fostering ownership of learning" sought to enhancing student engagement by exploring pedagogical solutions. Employing action research integrates action, evaluation, and reflection to drive classroom change. Findings reveal that student-centred practices like project-based learning foster increased interest, motivation, and active participation in learning experiences.

Alordiah's "Evaluation of a research training workshop for academic staff in tertiary institutions: A Kirkpatrick model approach" evaluates a workshop's impact on academic staff in Delta State, Nigeria, focusing on writing and publishing scholarly papers. Using the Kirkpatrick Model, results showed

high participant satisfaction and significant improvements in writing, publication, and research exposure, emphasising the workshop's effectiveness in enhancing research capabilities and academic recognition.

Le Pham et al.'s "Professional development activities of English language lecturers in Vietnam through the lens of sociocultural theory" explores professional development's impact on 56 English lecturers in Vietnam. Surveys and interviews reveal feedback, workshops, and resources like videos and websites crucial for their development. Social interaction and resources enhance their skills, knowledge, and motivation, suggesting regular training and resource availability for ongoing growth.

Wong and Chapman's "Development and validation of an instrument to measure expectancy for success and subjective task value constructs in the context of higher education" aimed to develop and validate an instrument in higher education. In surveying 565 undergraduate students from a large private institution in Singapore, two versions of the Expectancies and Values in Higher Education Instrument (EVHEI) were utilised. Results suggest that the EVHEI holds considerable promise for measuring motivation-related constructs at the higher education level.

Ó Murchú and O'Donoghue's "Advice from retired secondary school principals in Ireland on how to lead as a principal" aim to generate theory regarding the perspectives of recently retired secondary school principals in Ireland. The goal is to offer insights to providers, including university-led programmes, to inform the preparation of aspiring and appointed principals. The paper has four parts: rationale, recent developments in Irish school leadership, study methodology, and study results.

Ó Murchú and O'Donoghue's study concludes the research section and leads to two illuminating interview pieces. We start with an interview with an educational thought leader, Professor Rose Luckin, "Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin". Rose Luckin shares her journey into AI in Education (AIE), addressing gender bias and women's challenges. She also discusses other aspects, such as the ethical dimensions of AI deployment, advocating for learner-centred AI methodologies and stresses collaboration between educators and tech developers. In addition, Luckin evaluates generative AI's impact on assessment and learning in K-12 and higher education, emphasising lifelong learning and the need for collaborative efforts and ongoing research in navigating AIE's challenges and opportunities within ethical frameworks.

The interviews conclude with Brookfield et al.'s "'Failing well' in teaching about race, racism and white supremacy. An interview with Stephen Brookfield". The interview discusses his extensive international experience in education. It explores core concepts of race, racism, and white supremacy, reframing racism as systemic rather than individual. Brookfield emphasises an intersectional analysis, addressing racism in higher education and advocating for continuous antiracist efforts, challenging the idea of 'good white people' and promoting 'failing well' in the journey towards

antiracism. This interview is a much-extended version of a chapter in Brookfield et al.'s (2023) *Teaching well* and our third interview with Professor Stephen Brookfield in JALT (see Brookfield et al., 2019, 2022).

The Ed-tech review in this issue takes us back to the theme of AI. Perkins and Roe's "The use of Generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT" introduces a methodological innovation combining Generative AI (GenAI) tools with traditional qualitative research methodologies for thematic analysis. The approach enhances data processing and theme identification while maintaining the interpretative depth of human analysis. Challenges include managing inconsistencies in GenAI outputs and ensuring research validity through rigorous validation processes. The findings suggest a complementary relationship between GenAI and human researchers, accelerating analytical processes while leveraging human expertise and critical engagement.

The next section encompasses three opinion pieces, the first being Sam's "Reading between the lines: The necessity of books". This opinion piece celebrates the enduring significance of books in education, tracing their historical evolution and highlighting their role as repositories of human wisdom and culture. Amidst the digital age, books in physical form maintain their allure, offering tangible engagement and serving as lifelong companions in the pursuit of knowledge. In an era of technological advancement, the value of written knowledge in books remains timeless, guiding readers towards comprehension and enlightenment.

This is followed by Dey and Chakraborty's "Cargo cultism and the whiteness syndrome: Fake internationalization of private universities of India", which critiques the internationalisation efforts of Indian universities, identifying phenomena of cargo cultism and whiteness syndrome. Drawing from personal conversations with research participants, informal discussions with friends and colleagues, and analysing social media content, the piece challenges the effectiveness and authenticity of these initiatives.

This section concludes with Ifelebuegu's "Rise of the robots: What it means for educators" discussing an AI robot named Abigail Bailey as one of its "co-headteachers," illustrated the growing integration of AI in educational settings. This development prompts questions about the potential impact on traditional educator roles. This opinion piece explored the complexities of this issue, considering various factors that must be examined.

Cavagnari-Bruce et al.'s "Recognition of foreign professional degrees in Peru: Processes and strategies for improvement" falls into our category of a 'brief article'. It outlines the recognition process for foreign professional degrees in Peru and highlights the administrative nature of the recognition process and its lack of academic evaluation criteria. The authors suggest the need for academic assessments to ensure professionals meet high standards for entry into the job market.

Finally, we draw the curtains to a close for this issue with several book reviews. We start with Professor Waring's book review of Brookfield et al.'s (2023) *Teaching well: Understanding key dynamics of learning-centered classrooms. What does it mean to teach well?* Brookfield et al. (2023) delve into this question in *Teaching well*, exploring the essence of effective teaching through the lens of renowned scholar Stephen Brookfield. The book, co-authored by two of the authors of this editorial (Rudolph and Tan), examines key dynamics in learning-centered classrooms, covering topics such as classroom democratisation, critical thinking, and race. Each chapter poses pivotal questions to refine teaching practices and spark meaningful dialogue. With practical advice on integrating educators' identities into their pedagogy, it is Brookfield et al.'s hope to have provided an engaging text and a valuable resource for college and university educators worldwide.

This section concludes with Rudolph's book reviews of two gargantuan AI handbooks. Lindgren's (Ed., 2023) "Handbook of critical studies of artificial intelligence" comprehensively examines AI's societal impact, gathering insights from scholars worldwide. Lindgren's Handbook is a vital resource for academics, practitioners, and policymakers navigating AI's complex landscape, challenging prevailing techno-optimism with critical analysis and advocating for technology aligned with societal well-being. Du Boulay et al.'s (2023) "Handbook of artificial intelligence in education" thoroughly examines the field's development and practicalities, encompassing theories, methodologies, and future trajectories. Authored by esteemed scholars, its audience comprises researchers and advanced computer science, education, and AI students. Though the technicalities within the book may be a challenge for some readers, the Handbook's extensive coverage and insights render it a valuable asset for scholars and practitioners and a notable addition to AIED scholarship.

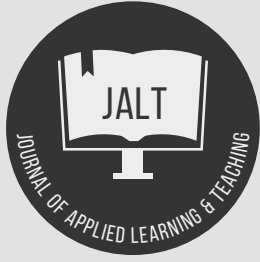
## References

- Andrew, M. B. (2024). The Great Resignation: The simple joys of not belonging. *Journal of Applied Learning & Teaching*, 7(1), 1-13. <https://doi.org/10.37074/jalt.2024.7.1.36>
- Barry, J. M. (2020). *The great influenza: The story of the deadliest pandemic in history*. Penguin.
- Benjamin, W. (2015). *Illuminations*. Bodley Head.
- Birmingham, K. (2014). *The most dangerous book: The battle for James Joyce's Ulysses*. Head Zeus.
- Bostrom, N. (2017). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Bridle, J. (2023). *New dark age: Technology and the end of the future*. Verso.
- Brookfield, S. D., Rudolph, J., & Tan, S. (2022). Powerful teaching, the paradox of empowerment and the powers of Foucault. An interview with Professor Stephen Brookfield. *Journal of Applied Learning and Teaching*, 5(1), 131-145. <https://doi.org/10.37074/jalt.2022.5.12>

- Brookfield, S. D., Rudolph, J., & Tan, S. (2023). *Teaching well*. Routledge.
- Brookfield, S. D., Rudolph, J., & Yeo, E. (2019). The power of critical thinking in learning and teaching. An interview with Professor Stephen D. Brookfield. *Journal of Applied Learning and Teaching*, 2(2), 76-90. <https://doi.org/10.37074/jalt.2019.2.2.11>
- Carr, N. (2020). *The shallows: How the Internet is changing the way we think, read and remember*. Atlantic Books.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Deming, R. (Ed.). (2013). *James Joyce. Volume I: 1907-27*. Routledge.
- Douglas, J. (1922, May 28). The beauty—and the beast. *Sunday Express*, p. 5.
- Eastman, M. (1931). *The literary mind: Its place in an age of science*. C. Scribners Sons.
- Eliot, T. S. (1975). *Selected Prose of T. S. Eliot*. F. Kermodé (Ed.). Faber and Faber.
- Ellman, R. (1982). *James Joyce. New and revised edition*. Oxford University Press.
- Freedom House. (2023). *Repressive power of artificial intelligence*. <https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence>
- Gardner, H. (1993). *Multiple intelligences: The theory in practice*. Basic Books.
- Gifford, D. (1988). *Ulysses annotated. Notes for James Joyce's Ulysses*. Revised and expanded edition. University of California Press.
- Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mädche, A., Röglinger, R., Ruiner, C., Schoch, M., Schoop, M., Urbach, N., & Vandirsk, S. (2023, March 20). *Unlocking the power of generative AI models and systems such as GPT-4 and ChatGPT for higher education: A guide for students and lecturers*. University of Hohenheim. <http://dx.doi.org/10.13140/RG.2.2.20710.09287/2>
- Goethe, J. W. (1997). *Faust. Erster und zweiter Teil* [Parts I & II]. Dtv.
- Goethe, J. W. (2003). *Faust. Parts I & II* (A. S. Kline, Trans). <https://antilogicalism.com/wp-content/uploads/2017/07/faust.pdf>
- Hammond, G. (2024, April 9). Elon Musk predicts AI will overtake human intelligence next year. *Financial Times*. <https://www.ft.com/content/027b133f-f7e3-459d-95bf-8afd815ae23d>
- Hastings, P. (2022). *The guide to James Joyce's Ulysses*. Johns Hopkins University Press.
- Homer. (2006). *The Odyssey*. (R. Fagles, Trans.). Penguin Classics.
- Hutchins, P. (2016). *James Joyce's world*. Routledge.
- Ismail, F., Crawford, J., Tan, S., Rudolph, J., Tan, E., Seah, P., ... & Kane, M. (2024a). Artificial intelligence in higher education database (AIHE V1): Introducing an open-access repository. *Journal of Applied Learning and Teaching*, 7(1), 1-9. <https://doi.org/10.37074/jalt.2024.7.1.35>
- Ismail, F., Crawford, J., Tan, S., Rudolph, J., Tan, E., Seah, P., ... & Kane, M. (2024b). Artificial intelligence in higher education database (AIHE V1). *Journal of Applied Learning and Teaching*, 7(1). <https://doi.org/10.37074/jalt.2024.7.1.35D>
- Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, 6(2), 56-63. <https://doi.org/10.37074/jalt.2023.6.2.34>
- Joyce, J. (2000a). *Ulysses*. Penguin Classics.
- Joyce, J. (2000b). *A portrait of the artist as a young man*. Penguin Classics.
- Joyce, J. (2008a). *Ulysses. The Gabler Edition*. Random House.
- Joyce, J. (2008b). *Dubliners*. Oxford University Press.
- Jung, C. G. (1979). *The collected works of CG Jung: The spirit of man, art and literature (Vol. 15)*. Routledge and K. Paul.
- Knox, B. (2006). Introduction. In Homer, *The Odyssey* (pp. 3-64). Penguin Classics.
- Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Penguin.
- Lawrence, D. H. (2006). *Lady Chatterley's lover*. Penguin Classics.
- Lindgren, S. (2023a). *Critical theory of AI*. Polity.
- Lindgren, S. (2023b). Introduction. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 1-19). Edward Elgar.
- Luckin, R., Rudolph, J., Grünert, M., & Tan, S. (2024). Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin. *Journal of Applied Learning and Teaching*, 7(1), 1-18. <https://doi.org/10.37074/jalt.2024.7.1.27>
- Mais, S. B. P. (1922, March 25). An Irish revel. *The Daily Express*.
- Marr, B. (2023, May 11). The hot new job that pays six figures: AI prompt engineering. *Forbes*. <https://www.forbes.com/sites/bernardmarr/2023/05/11/the-hot-new-job-that-pays-six-figures-ai-prompt-engineering/?sh=63f02a1a7d7f>

- Mchangama, J., & White, J. (2024, February 26). The future of censorship is AI-generated. *Time*. <https://time.com/6835213/the-future-of-censorship-is-ai-generated/>
- McNamara, M. (2010, May 25). Television review: The '24' finale. *Los Angeles Times*. <https://www.latimes.com/archives/la-xpm-2010-may-25-la-et-0525-24-review-20100525-story.html>
- Mills, A. (2023). ChatGPT just got better. What does that mean for our writing assignments? *Chronicle of Higher Education*. <https://www.chronicle.com/article/chatgpt-just-got-better-what-does-that-mean-for-our-writing-assignments>
- Nabokov, V. (1990). *Strong opinions*. Random House.
- Pinker, S. (2014). *The sense of style. The thinking person's guide to writing in the 21st century*. Penguin.
- Pomerantsev, P. (2017). *Nothing is true and everything is possible: Adventures in modern Russia*. Penguin.
- Popenici, S. (2023). *Artificial intelligence and learning futures. Critical narratives of technology and imagination in higher education*. Routledge.
- Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023). A critical perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning and Teaching*, 6(2), 311-331. <https://doi.org/10.37074/jalt.2023.6.2.5>
- Potter, R. (2009). Obscene modernism and the trade in salacious books. *Modernism/modernity*, 16(1), 87-104. <https://doi.org/10.1353/mod.0.0065>
- Preskill, S., & Brookfield, S. D. (2008). *Learning as a way of leading: Lessons from the struggle for social justice*. John Wiley & Sons.
- Rudolph, J., Crawford, J., Sam, C. Y., & Tan, S. (2024a). Introduction. In J. Rudolph, J. Crawford, C. Y. Sam, & S. Tan (Eds.), *Handbook of crisis leadership in higher education*. Ahead of print.
- Rudolph, J., Ismail, M. F., & Popenici, S. (2024b). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. *Journal of University Teaching and Learning Practice*, 21(6). <https://doi.org/10.53761/54fs5e77>
- Rudolph, J., & Tan, S. (2022). The war in Ukraine as an opportunity to teach critical thinking. *Journal of Applied Learning and Teaching*, 5(1), 165-173. <https://doi.org/10.37074/jalt.2022.5.1.24>
- Rudolph, J., Tan, S., & Aspland, T. (2022). Editorial 5(2): Avoiding Faustian pacts: Beyond despair, impostorship, and conceit. *Journal of Applied Learning and Teaching*, 5(2), 06-13. <https://doi.org/10.37074/jalt.2022.5.2.1>
- Rudolph, J., Tan, S., & Tan, S. (2023a). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Rudolph, J., Tan, S., & Tan, S. (2023b). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Sam, C. Y. (2024). Reading between the lines: The necessity of books. *Journal of Applied Learning and Teaching*, 7(1), 1-10. <https://doi.org/10.37074/jalt.2024.7.1.2>
- Tan, S., Rudolph, J., & Tan, S. (2024). Riding the generative AI tsunami: Addressing the teaching and learning crisis in higher education. In J. Rudolph, J. Crawford, C. Y. Sam, & S. Tan (Eds.), *Handbook of crisis leadership in higher education*. Ahead of print.
- Verdicchio, M. (2023). Marking the lines of artificial intelligence. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 245-253). Edward Elgar. <https://doi.org/10.4337/9781803928562.00027>
- Woolf, V. (1923, April 5). How it strikes a contemporary. *The Times Literary Supplement*.
- Woolf, V. (1980). *The Diary of Virginia Woolf, Volume 2: 1920-1924*. Harvest Book.
- Woolf, V. (2018). *Mrs. Dalloway*. Penguin Classics.
- World Economic Forum. (2024). *Global risks report 2024: Digest*. <https://www.weforum.org/publications/global-risks-report-2024/digest/>

Copyright: © 2024. Jürgen Rudolph, Shannon Tan and Fadhil Mohamed bin Mohamed Ismail. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



---

## Artificial intelligence and graduate employability: What should we teach Generation AI?

---

Peter Waring<sup>A</sup>

A

*Professor, Pro Vice Chancellor of Transnational Education and Dean, Murdoch University Singapore*

DOI: <https://doi.org/10.37074/jalt.2024.7.1.42>

---

### Abstract

In the era of Generative Artificial Intelligence (AI), universities are grappling with how best to prepare students for a workforce increasingly influenced by the technology. This commentary explores the implications of AI for graduate employability, emphasising the need for educational institutions to reassess their curricula. It suggests integrating AI literacy and ethical decision-making skills to ensure graduates remain valued by employers. As AI encroaches on graduate-entry and even high-skill jobs, I call for a curriculum that balances technical skills with critical thinking and interpersonal abilities and that prepares students for the complexities of a technology-driven job market.

**Keywords:** AI; artificial intelligence; ChatGPT; future of work; generative artificial intelligence; graduate employability; LAMs; Large Action Models; Large Language Models; LLMs.

The defeat of Chess Grandmaster Gary Kasparov to IBM's Deep Blue computer in a game of chess in 1997 now seems like a pivotal moment in the history of machine intelligence. It was the first public display of artificial intelligence's (AI) capacity to overtake the very best human performance, albeit in the narrow arena of chess. This perhaps should have been a prescient moment that foreshadowed the emergence of Generative AI, but in truth, much of the higher education sector was caught off-guard by the explosive adoption of AI tools such as ChatGPT in 2022 (Rudolph et al., 2023b).

OpenAI's ChatGPT sudden entrance on the world stage sparked immediate concerns throughout the global higher education sector. Chief among these was the academic integrity implications of Large Language Models (LLMs). How could universities ensure the authenticity of assessment, if essays, reports, coding tasks and term papers could be so easily manufactured by Generative AI tools in mere moments? A rigorous discussion ensued across the sector with various solutions suggested from returning to high-value invigilated exams through to assessment re-design and increased use of oral examination (*viva voce*). Higher Education regulators such as Australia's TEQSA responded

by developing whole repositories of informed responses to AI and suggested guidelines for learning and teaching (see TEQSA, 2024).

The conundrum faced by educators as a result of AI extends beyond authenticity and trust in assessment. It also poses serious and challenging questions as to what universities should teach, how they should teach and how they should best prepare students for the rapidly changing world of work. While universities tend to have multiple and diverse missions, it is generally accepted that ensuring graduates are employable and considered valuable to prospective employers is a common aspiration. The evolution of AI is occurring at unprecedented speed, spurred on by both rapid technological developments in AI chip design and remarkable investor support for AI applications from everything from electric vehicles to smartphones, airports, and green technologies. In this opinion piece, I explore the challenge to graduate employability posed by AI models and tools and the question it raises for universities as to how best to prepare the next generation (Generation AI) for employment likely to be heavily laden with the influence of AI.

Generative AI, to some extent, has upended previous predictions of the future of work and autonomous systems. Prior to ChatGPT, there was a strong view in the 'Future of Work' literature that, like a rising tide, machine intelligence and autonomous systems would replace tasks requiring lower skills and cognitive ability (see Waring et al., 2020). Higher-level functions requiring advanced cognitive ability – for instance, creativity, reasoning, problem-solving, literacy, numeracy, judgment, translation and interpretation, and emotional and interpersonal skills – were predicted to be less vulnerable to the 'great replacement machine' of AI. With perhaps a touch of hindsight, the 'Future of Work' literature has sometimes appeared analogous to debates regarding Science and Creationism and the origin of the Universe. In that age-old discussion, it is Creationism that with each new scientific discovery (for example, that the Earth is not the centre of the Universe) makes certain accommodations. Similarly, much of the 'Future of Work' literature has previously claimed that jobs of the future will need to emphasise those qualities that are innately

human – interpersonal skills and emotional intelligence are often cited as among these human qualities that cannot be replaced with silicon-based forms of intelligence. That may be so presently, but time and the march of technology are proving this to be wishful thinking. Since the rapid arrival of Generative AI, the world has witnessed the credible emulation of a range of human abilities at super-human speed.

The steady encroachment of AI towards emulating what was previously thought of as innately and unreproducible human qualities has been a feature of its evolution thus far. Multi-modal AI models such as Google DeepMind's Gemini are already able to reason across a range of inputs, including audio, video, text, code and images. Google also claims that Gemini outperforms human experts on 'massive multitask language understanding' (Pykes, 2024).

There are also Large Action Models (LAM) emerging, which not only perform the same range of tasks as LLMs but also understand and perform actions that would otherwise be undertaken by humans. A recent demonstration of this technology by firm Rabbit Tech demonstrated how their AI tool could not only propose a complex travel itinerary that incorporated specific spoken wishes but also proceed with making all the necessary travel bookings (Pan, 2024). It is highly probable that these LAM versions of AI will further evolve and become more adept at undertaking a range of actions at the instruction of human beings or with some level of autonomy.

Many of the jobs that require higher-level cognitive capabilities are professional, white-collar jobs – and many of these careers begin at graduate entry. As Cazzaniga et al. (2024) note, these AI tools and models challenge the belief that technology affects mainly middle and, in some cases, low-skill jobs; its advanced algorithms can now augment or replace high-skill roles previously thought immune to automation. For the creators of these AI tools, there is an economic incentive tied to expanding the range of functions and utility of AI. Thus, it might be expected that AI will continue to climb the value chain and replace the higher-level tasks characteristic of well-paid jobs. This rather surprising trajectory of AI technology illustrates the difficulty in making reliable predictions when it comes to new technologies, thus rendering the traditional educative role of universities increasingly uncertain. In strategic terms, the concept of 'skating to where the puck is going to be' (often attributed to ice hockey legend Wayne Gretzky), which denotes getting ahead of the strategic curve, has become extremely challenging for university leaders.

It is also reasonably foreseeable that in the future, professions that require high-level empathy, compassion, and complex interpersonal skills (Psychology, Counselling, Nursing, etc.) could be substantially replaced by AI bots with advanced 'empathy algorithms' – able to listen compassionately and offer curated advice based on many terabytes of clinical psychology data. Already customer contact or call centres are looking to replace human beings with AI bots that can listen to and manage customer enquiries with all the natural language processing and interpersonal skills of a call centre worker (Valentino, 2024).

Adding to this uncertainty are the variable effects of AI across jurisdictions and industries. Cazzaniga et al. (2024) note that the impact of AI is likely to be uneven across different sectors and across different countries. In particular, their study demonstrates that 60 per cent of jobs in high-income/advanced countries are exposed to AI due to "the prevalence of cognitive-task orientated jobs" (p. 2). By contrast, the proportion of AI-exposed jobs in less developed countries is estimated to be 26 per cent. Yet they argue that this does not mean that workers in high-income countries will be necessarily worse off. Instead, Cazzaniga et al. (2024) claim that AI is more likely to complement rather than displace the employment of high-income workers. This is because they speculate that AI tools are most likely to make university-educated workers more productive. Furthermore, it is argued that the productivity gains from AI penetration are expected to boost total income in high-income countries. It is important, however, to caution that Cazzaniga et al. (2024) suggest that the adoption of AI is likely to amplify income and wealth inequality if the productivity gains are captured by a minority of those who own the technology, have equity stakes in AI companies or who are skilled in using the technology.

Consistent with Cazzaniga et al. (2024), the World Economic Forum's (2023) research (covering 803 companies employing 11.3 million workers) on the 'Future of Jobs' also acknowledges the prospect of 'job destruction' from AI and automation. However, the 'Future of Jobs' report also expects that the overall impact of technological change will be 'net positive' for job growth based on their employer survey. Interestingly, the WEF survey finds (perhaps counter-intuitively) that businesses expect to introduce automation at a slower pace than previously believed. Overall, they find that while the human-machine frontier is shifting in favour of tasks performed by machines, they estimate that two-thirds of all tasks are still performed by humans (World Economic Forum, 2023). This may suggest that many businesses are just beginning to examine the application of AI tools.

Universities, though, should be concerned for the future of graduate-entry positions. Many of these roles across the functional areas of businesses tend to be structured around tasks requiring less advanced cognitive abilities that are also computer-based. Tasks such as researching a topic, writing emails or short reports, organising meetings, constructing presentation decks, performing spreadsheet calculations, analysing financial statements, drafting contracts, and writing simple computer code are typical of a range of graduate-entry positions (see Rudolph et al., 2023a). These are exactly the kinds of tasks that are clearly within the capability set of Large Language Models, performed at digital speeds at near zero cost. As Brown (2023) notes, the adoption of Generative AI by organisations threatens to remove 'the bottom rung of the ladder' for those graduates starting their careers. This is persuasive and points to the need for graduates to be 'AI literate' so that they not only adroitly use the AI tools they need to complete tasks and solve problems but also have the ability to critically assess the output of AI models.

There is potentially another, more subtle and negative impact on graduate jobs, which is rarely acknowledged in the growing AI literature. That is the experience that performing

these less complex tasks brings, and the subtle qualities of discipline, focus and persistence that are ingrained in graduates through performing routine tasks. These are presumably lost (or need to be developed in other ways) when AI is routinely used to perform these tasks.

Where does all this leave university graduates and what can universities do to ensure that graduates remain valued by employers?

Brown (2023) calls for universities to urgently review their employability strategies to account for the growth in AI while also integrating AI into teaching and learning. Importantly, he builds a case for universities to emphasise the development of “[c]ritical analysis, critical enquiry, problem formulation, socio-ethical considerations, interpersonal skills, resilience” as key graduate attributes in a world with near-universal adoption of AI (p. 20). Brown (2023, p. 19) stresses that the use of AI is “likely to place a premium on critical thinking skills, including the ability to challenge and interrogate knowledge”. Further, he contends that its widespread adoption will increase the value attached to interpersonal skills.

Similarly, ‘The Future of Jobs’ research published by the World Economic Forum indicated that employers regarded ‘Analytical Skills’ as being the most important core skill. ‘Creative Thinking’ ranks second, but interestingly, self-efficacy skills such as resilience, flexibility, agility, motivation, self-awareness, curiosity, and lifelong learning are also highly ranked, which suggests that employers suspect that these qualities will prove important for technology-disrupted workplaces (World Economic Forum, 2023). In the same study, ‘AI and Big Data’ skills saw the biggest increase in ranking by employers, signalling that they expect to invest significantly in upskilling their workforces to be able to use AI tools effectively. This was especially true of large employers (those with more than 50,000 employees) who responded to the WEF survey.

The importance of ‘self-awareness’ as a core skill to be developed in the age of AI is also advanced by historian and public intellectual Yuval Noah Harari who argues that in the face of relentless change, people will need to constantly reinvent themselves. Achieving this requires mental flexibility, resilience, and emotional intelligence (Irais, 2023). Harari argues, “Investing in people’s flexibility and mental or psychological resilience is no longer a luxury. It’s essential to survive in the 21st Century job market” (Irais, 2023). Perhaps at a more philosophical level, Harari makes the point that technology provides the human race with unprecedented power. Therefore, there is a need to teach people how to exercise that power ethically and responsibly. Part of this will also require teaching people to make good decisions, to be able to critically evaluate a set of circumstances or fact patterns to determine what is fact/evidence and what is not – to judge what is reliable information and what is unreliable.

I now propose an input/output and action model to inform how universities can come to grips with designing a curriculum to meet the challenges of the AI generation. The challenge for higher education institutions in designing

a curriculum for the AI era is considerable. The shifting task frontier between humans and machines injects considerable uncertainty into curriculum design. If history is quintessentially the study of change, as Harari has argued, then we are witnessing history unfold at a rapid rate as a consequence of unrelenting technological innovation.

As I have already contended, there is an urgent need for universities and other institutes of higher learning to develop curricula that equip graduates with a set of AI competencies to make effective use of the new technology. Part of this will require educators to explain how AI models work – to explain how they are designed, built and trained. But another, perhaps more significant, need is for educators to teach students how to ethically make the greatest use of these tools.

To this end I would like to suggest that educators think of this challenge in terms of an ‘Input/Output/Action’ model. The model that I advance here, deconstructs the processes by which AI tools are used and identifies a set of questions at each stage that should inform the development of curricula and graduate competencies.

The ‘Input’ stage recognises that AI tools require some level of input that typically comes from the human user of the tool. This might be a question or prompt, an image, a video file or perhaps computer code. At this initial stage, it is relevant to ask about the type and quality of information that is submitted to the AI tool. How do users ensure that the input is relevant, reliable and, if a question, framed in a way that it is likely to produce the best results? This also requires an appreciation of the data sets on which AI models are developed and the possibility that these large data sets may contain errors and bias. Educators, therefore, need to develop the critical faculties of students who are using these tools to be able to ask these questions and understand the limitations of AI data sets and the methods of ‘learning’.

At the ‘Output’ stage, there is an equally important need to be able to interrogate what is produced by the AI tool. Users need to be able to determine if the outputs are valid, reliable, relevant, and grounded in reality. The AI literature (see Naddaf, 2023; Zhou et al., 2023) has demonstrated quite consistently that AI tools are capable of fabricating output which is entirely fictional, including, for example (and worryingly), in areas of medical science and the law. To be able to discern if the AI output is valid and not the product of an AI ‘hallucination’, users need to have some subject knowledge or know how to check the veracity of the AI output (see Rudolph et al., 2023b). Additionally, users need to be able to assess if the AI output is potentially biased or offering unethical or ethically dubious advice. Thus, it will be important for educators to teach students how to think ethically and apply ethical principles to AI output.

The final stage in the model draws attention to the set of skills required to action the output of an AI tool. It is one thing to generate AI output but quite another to use that output effectively and responsibly. University graduates will need to know what actions to take based on the output of the AI. This implies that curriculum and training on ethical decision-making will be important for those charged with



actioning AI output. Effectively executing the advice or output of AI tools will also require graduates to have good collaboration, communication, leadership negotiation, intercultural and teamwork skills. Getting things done might be expedited through the use of AI, but subsequent decisions, collaborations, and actions depend on a set of technical and interpersonal skills and knowledge that are competently exercised. Further, users of AI tools will also need to know how to evaluate the results of the actions they have taken and to reflect on how future actions could be improved based on experience.

#### Input

- Do we understand the task/problem properly?
- Is AI the most effective means to address the problem?
- Are we potentially sharing sensitive information or Intellectual Property by using AI?
- Which AI tool/model is best to explore our question?
- How do we frame the right question/prompt?
- Is our question subject to implicit biases or misperceptions?

#### Output

- What is the nature of the output and does it seem to align with the question asked?
- How do we know if the output is valid and reliable? By what standard should we assess the information?
- Does the information present as biased or the result of discriminatory reasoning?
- Is the information consistent with ethical principles and reasoning?

#### Action

- What action might be appropriate based on the output?
- Have we consulted with the right stakeholders concerning actions that might be taken?
- If the action involves 'change' have we considered the social and psychological impact of the change? What is the most effective, least damaging way of implementing the change?
- Have we considered consequences of acting on the output?

Figure 1. Proposed Input-Output-Action Model.

This simple Input/Output/Action model serves to highlight a set of competencies that are proposed for using AI responsibly and effectively. Universities, in my opinion, need to ensure that their curriculum and student learning experience develop these skills and knowledge. As Cazzaniga et al. (2024) claim, those who can skilfully use AI are more likely to be valued by employers and enjoy higher incomes as a consequence. Although learning, particularly at a tertiary level, is not simply about gaining meaningful and well-rewarded employment, it is undeniably a significant part of the mission of universities to prepare their graduates as best they can for an ever-changing world.

## References

Brown, R. (2023, November) *The AI Generation: How universities can prepare students for the changing world*. Demos. <https://demos.co.uk/wp-content/uploads/2023/11/The-AI-Generation-2.pdf>

Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A., Pizzinelli, C., Rockall, E., & Taveres, M. (2024, January 14). *GEN-AI: Artificial Intelligence and the future of work, staff discussion note*. International Monetary Fund.

Irais, S. (2023, October 19). *Flexible minds, a challenge for the future of education: Yuval Harari*. Conecta, [www.conecta.tec.mx](http://www.conecta.tec.mx)

Naddaf, M (2023, November 22) 'ChatGPT generates fake data set to support scientific hypothesis'. Nature News, [www.nature.com](http://www.nature.com).

Pan, C. (2024, January 22). *Rabbit, US\$200 AI-powered gadget, becomes unexpected hit after CES 2024 launch by Chinese entrepreneur*. South China Morning Post.

Pykes, K. (2024, January). *What is Google Gemini? Everything you need to know about Google's ChatGPT rival*. DataCamp. <https://www.datacamp.com/tutorial/what-is-google-gemini>

QS (2021). *How Artificial Intelligence is influencing graduate employability and the global higher education sector*. [www.qs.com](http://www.qs.com).

Rudolph, J., Tan, S., & Aspland, T. (2023a). Editorial 6(1): Fully automated luxury communism or Turing Trap? Graduate employability in the generative AI Age. *Journal of Applied Learning & Teaching*, 6(1), 7-15. <https://doi.org/10.37074/jalt.2023.6.1.35>

Rudolph, J, Tan, S., & Tan, S. (2023b) ChatGPT: Bullshit spewer or the end of traditional assessments in higher education', *Journal of Applied Learning & Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>

TEQSA. (2024). *Higher education good practice hub – Artificial intelligence*. [www.teqsa.gov.au](http://www.teqsa.gov.au)

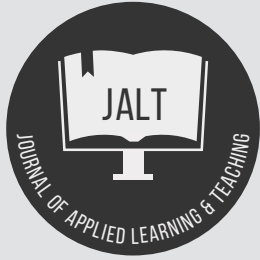
Valentino, A. (2024, April 23). *Will AI customer service kill the call centre?* TechMonitor. <https://techmonitor.ai/technology/ai-and-automation/ai-customer-service-llm>

Waring, P., Bali, A., & Vas, C. (2020). The fourth industrial revolution and labour market regulation in Singapore. *The Economic and Labour Relations Review*, 31(3), 347-363. <https://doi.org/10.1177/1035304620941272>

World Economic Forum. (2023). *Future of jobs report 2023*. <https://www.weforum.org/publications/the-future-of-jobs-report-2023/>

Zhou, J., Zhang, Y., Luo, Q., Parker, A., & Choudhury, M. (2023, April). Synthetic lies: Understanding AI-generated misinformation and evaluating algorithmic and human solutions. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany. <https://doi.org/10.1145/3544548.3581318>

Copyright: © 2024. Peter Waring. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Thematic exploration of educational research after the COVID pandemic through topic modelling

---

Indu Bala<sup>A</sup>

A

*School of Computer and Mathematical Sciences, The University of Adelaide, Adelaide, Australia, 5005*

Lewis Mitchell<sup>B</sup>

B

*School of Computer and Mathematical Sciences, The University of Adelaide, Adelaide, Australia, 5005*

---

### Keywords

Academic journal trends;  
academic teaching;  
BERTopic model;  
educational research;  
higher education;  
learning analytics.

---

### Abstract

The Journal of Applied Learning and Teaching is known for its focus on innovative practices in learning and teaching in higher education. In this study, we utilised BERTopic modelling to investigate trends and research within the journal. Our objective was to analyse thematic structures and identify emerging trends in a vast academic research corpus. BERTopic modelling enabled us to categorise academic texts into distinct topics, revealing underlying patterns and themes. Our analysis unveiled various topics, showcasing the journal's interdisciplinary nature. Particularly, articles from January 2021 to December 2023 shed light on global trends in learning and teaching amidst significant changes in the post-COVID era. We identified 17 frequent topics, categorised into four major thematic groups: Technology and Digital Learning in Education, Healthcare and Clinical Training, Educational Strategies and Outcomes, and Pandemic-Driven Social and Compassion Aspects in Education. We examined these themes and presented the findings, highlighting challenges and opportunities in higher education. This comprehensive analysis serves as a roadmap for future research, guiding scholars and practitioners in advancing applied learning and teaching.

---

### Correspondence

[indu.bala@adelaide.edu.au](mailto:indu.bala@adelaide.edu.au)<sup>A</sup>

---

### Article Info

Received 20 February 2024

Received in revised form 21 March 2024

Accepted 22 April 2024

Available online 23 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.40>

## Introduction

The education landscape is continually evolving, driven by technological advancements and global challenges. Innovative teaching methods and learning paradigms are necessary to maintain a positive learning environment (Escueta et al., 2017). Global challenges, such as COVID-19 and wars, can sometimes disrupt or alter the learning environment. Therefore, it is crucial to periodically evaluate the existing methods in light of current situations and prepare for the future. In this context, understanding the thematic progression in higher educational research is vital (Yunita, 2018).

In recent years, there has been a surge in research aiming to understand the thematic progression of educational landscapes, resulting in a significant increase in the application of artificial intelligence in academic research (Ai, 2017). This has introduced new methods for analysing scholarly literature. Additionally, there has been a growing demand for interdisciplinary approaches among scientific disciplines, leading to the development of numerous interdisciplinary programmes to address academic challenges. The outcomes of these programmes are annually published in various journals (Huston et al., 2018). However, it is essential to periodically analyse these research findings to comprehend current challenges in academia and gain insights for the future. Analysing the vast amount of published research is daunting due to the rapid advancement of science and technology. This underscores the importance of monitoring research trends to identify potential innovations. Research trends can be identified using various sources, including scientific literature, books, articles, and publications extensively reviewed by researchers worldwide (Ranjbar-Sahraei & Negenborn, 2017; Jiang et al., 2018). The analysis of these publications has proven helpful in identifying emerging topics and tracking their evolution over time. Topic modelling techniques offer one way to analyse these text-based published research articles, helping to uncover patterns and address more specific research questions (Amado et al., 2018).

Topic modelling has been effectively used to discern patterns and topics from scholarly publications. Researchers have previously endeavoured to determine trends through publications to analyse current and future directions. For example, Chen et al. (2020) utilised structural topic modelling to analyse articles published in the *Computers & Education Journal*, identifying research hotspots. They also analysed annual topic proportion trends and topic correlations, offering insights into potential future research directions within the journal's scope. Similarly, Pandur et al. (2020) employed a combined approach of Structural Topic Modelling and Latent Dirichlet Allocation (LDA) to extract topics and identify current trends from scientific papers in the field of social science from the Web of Science. Later, Lemay et al. (2021) conducted a structural topic modelling of articles in educational data mining and learning analytics, revealing thematic features of these two fields. They identified five significant topics within educational data mining and learning analytics and analysed the differences in research focus between the two disciplines.

Furthermore, Nylander et al. (2022) applied topic modelling to publications in the *International Journal of Lifelong Education*, identifying predominant themes and examining the evolution of the journal's content over time. Hussain et al. (2022) crafted a multi-layered topic modelling approach that integrates situation awareness with an advanced hybrid machine learning technique to analyse students' textual feedback in academic environments. Maphosa and Maphosa (2023) employed the LDA for a bibliometric analysis on a subset of the Scopus database, explicitly examining the progression of Artificial Intelligence (AI) research in higher education (HE). Choi and Lee (2023) constructed a topic map covering areas such as biocompatible materials, structural materials, electrochemistry, and photonics, using it to discern national research priorities in materials science and to explore the competitive stances and strategic approaches of leading countries. Additionally, in recognition of COVID-19's extensive impact across research fields, Cao et al. (2023) utilised topic modelling on published abstracts to evaluate the pandemic's effect on research directions. By applying the LDA method, they delineated the research topics, trends, and topic correlations in COVID-19 studies, finding that the thematic similarity between topics increased with the scope of documents analysed.

Focusing on trends in higher education post-COVID-19 pandemic, this study aims to analyse a collection of journal abstracts from the *Journal of Applied Learning and Teaching (JALT)*. JALT serves as a crucial hub, offering a global platform for new ideas and insights in the realm of higher education. Employing the sophisticated BERTopic modelling technique, this research identifies the most prominent themes and patterns that have emerged in recent years. This illuminates the evolving nature of academic research, especially in education and teaching, after the COVID-19 pandemic. Our study reveals detailed trends and shifts in higher education methodologies. The objectives of this study are:

- To conduct a BERTopic modelling-based analysis of the *Journal of Applied Learning and Teaching*, pinpointing the major research topics that have arisen following the COVID-19 pandemic.
- To provide a broader range of thematic research areas and analyse their annual distribution, extracting significant insights.

## Methodology

Text mining techniques are designed to unearth valuable knowledge that may be hidden or not immediately apparent within a vast amount of textual data. Standard text mining methods include unsupervised and supervised techniques such as text categorisation, text clustering, document summarisation, and keyword extraction (Gurcan & Cagiltay, 2023; Bala, 2023). Topic modelling, a subset of text mining, is an unsupervised machine-learning technique that identifies topics within a collection of documents. For our analysis, we utilised the BERTopic model, an advanced method that typically does not require extensive data preprocessing. However, given the complexity of our dataset and our goal to extract meaningful topics accurately, we performed a thorough data preprocessing, which we detail in the

subsequent sections.

## Data source and preprocessing

We extracted data from published research articles on the Journal of Applied Teaching and Learning (JALT) website from January 2021 to December 2023. The dataset comprises the title, research type, publication date, DOI link, abstract, and keywords for 144 articles. To pre-process the text from the abstract collection, we implemented the following steps:

- Lowercasing: All text data is converted to lowercase to standardise the text, ensuring that words in different cases (e.g., "Abstract", "abstract", "ABSTRACT") are treated as the same (Alasadi & Bhaya, 2017).
- Removing punctuation and numbers: Using regular expressions, the text is stripped of punctuation and numbers, identifying and removing non-word characters and numerical digits. This refinement focuses the analysis on the lexical content, removing extraneous elements that could skew the textual analysis.
- Tokenising the text: The cleaned text is tokenised and split into individual words or tokens using the word tokenise function from the Natural Language Toolkit (NLTK). This step breaks the text into words, provides a list of tokens, and sets the stage for further processing, such as stemming or lemmatisation (Hardeniya et al., 2016).
- Removing stopwords: Common words that typically carry little meaning (stopwords), such as 'the', 'is', 'in', etc., are removed using the NLTK-provided list of English stopwords. Filtering out these words eliminates those not likely to be significant for the analysis.
- Re-constructing the text: The remaining words (tokens) are then joined to form the final pre-processed text, which contains only the relevant and meaningful words ready for analysis.

## Topic modelling

BERTopic is a state-of-the-art topic modelling method in Natural Language Processing (NLP), employing transformer embeddings and clustering algorithms to extract meaningful topics from text (Grootendorst, 2022). The BERTopic process involves several key steps:

- Embedding of documents: Documents are converted into numerical formats using the sentence-transformer model "Paraphrase-multilingual-MiniLM-L12-v2". This pre-trained model effectively generates sentence embeddings and facilitates semantic search (Reimers & Gurevych, 2019).

- Reducing of dimensionality: BERTopic reduces the dimensionality of these embeddings using the Uniform Manifold Approximation and Projection (UMAP) technique, maintaining the integrity of local and global structures in reduced dimensions (McInnes et al., 2018).
- Clustering process: The core of topic extraction involves clustering the reduced embeddings into related groups using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), which excels at identifying clusters of varying densities and is resilient to noise (McInnes et al., 2017).
- Topic tokenisation: CountVectorizer, which converts text data into a numerical format, is used for tokenising topics, facilitating efficient analysis and interpretation of the results (Hu & Zhang, 2022).
- Weighting and Topic Differentiation: A class-based TF-IDF (c-TF-IDF) approach, which focuses on clusters rather than individual documents, is used to differentiate clusters. This method identifies unique aspects of documents within each cluster, treating each cluster as a single document and analysing word frequencies. This approach precisely defines clusters and suggests topic names (Grootendorst, 2022; Bala et al., 2023). Figure 1 visualises the working process of the BERTopic model.

BERTopic provides a flexible modelling approach that does not require the pre-definition of the number of topics. However, setting a potential topic range can help guide the model to capture broader themes (Wang et al., 2023; Bala et al., 2024). After training the model, it is crucial to evaluate the topics to ensure they are coherent, interpretable, and accurately reflect the main themes of the text data. This evaluation often involves assessing the coherence and distinctiveness of the topics. To further refine the results, topics can be merged based on similarity. This process entails setting a similarity threshold score; topics with a score above this threshold are grouped. This reduces the granularity of the topics and allows for broader thematic categories, making them easier to analyse and interpret.

In this study, we enhanced our analysis by setting a high similarity threshold of 0.83. Applying this to the 17 topics obtained from the BERTopic model, we successfully grouped them into four distinct thematic categories. This consolidation provided a clearer, more thematic organisation of the topics, aiding in a more straightforward understanding and interpretation of the subject matter in the corpus.

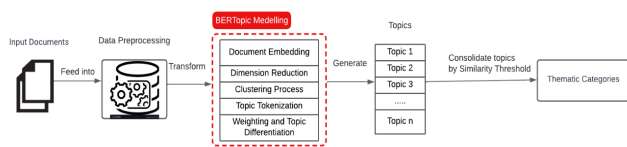


Figure 1. Graphical representation of BERTopic model.

## Results

### Data analysis

For the experimental evaluation, the data were analysed, and a notable observation was that a majority of the articles incorporated references to COVID, varying from minor mentions to significant discussions. Figure 2(a) displays the year-wise number of publications in the visual representation, indicating a consistent annual increase in published articles. Complementing this, Figure 2(b) offers a histogram depicting the word count distribution across the abstracts, highlighted by a distribution curve. The most frequent abstract length hovers around the 200-word mark, a crucial detail for text analysis considerations.

We pre-processed the dataset using Python packages before conducting the topic modelling analysis on abstracts. This preprocessing resulted in a total word count of 31,887 in the abstracts, which was the foundation for our topic modelling exercise. This preparatory step ensured that the subsequent analysis was conducted on a refined dataset, poised to yield more accurate and meaningful insights into the prevalent topics of academic discourse in higher education.

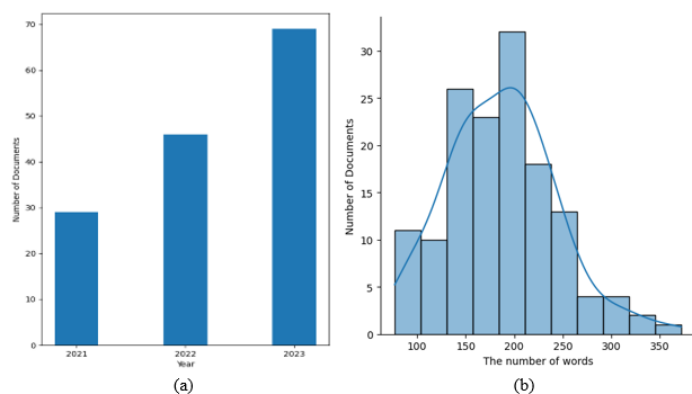


Figure 2. (a) Number of publications per year, (b) Word distributions of documents.

### Topic modelling results analysis

In this study, we applied the BERTopic model to a collection of abstracts, identifying 17 distinct topics. These were visually presented in Figure 3 through a bar graph that illustrates the distribution of topics across the dataset. Each bar's height indicates the prevalence or score of a topic, accompanied by the top 15 high-scoring words associated with these topics. This facilitated the assignment of descriptive labels to each topic, derived from the most significant and recurring terms that reflect their central themes. The highest-scoring terms within each topic were pivotal in the labelling process, as they most represent the topic's content.

Delving into the specifics, Topic 0, 'AI in Education,' features terms like 'AI,' 'ChatGPT,' and 'education,' suggesting a focus on AI tools such as chatbots in educational settings. Meanwhile, Topic 1, 'COVID-19 and Education,' includes words like 'COVID,' 'pandemic,' and 'online,' indicating discussions about the pandemic's impact on educational practices, particularly the shift to online learning. Topic 2, 'Online Learning Environments,' characterised by terms such as 'online,' 'learning,' and 'social,' points to the evolving nature of digital learning spaces and tools.

Further topics range from 'Educational Feedback and Outcomes,' examining the efficacy of feedback in learning, to 'Mathematics Education in Sub-Saharan Africa,' highlighting unique regional challenges and teaching methodologies. 'Digital Learning Platforms and Practices' underscores the growing trend in digital education, while 'Integrated Programming and Engineering Equation' and 'Teacher Professional Development and Learning' highlight the expanding scope of interdisciplinary educational approaches and continuous educator development.

In healthcare education, 'Clinical Skills and Healthcare Training' emphasises the importance of practical training. At the same time, 'Academic Connectedness and Resilience' highlights the significance of mental resilience and social bonds in academic settings. The theme of collaboration is further explored in 'Collaborative Online Teaching and Learning,' showcasing the value of community and interaction in online education. 'Graphic Design and Modern Pedagogical Approaches in Higher Education' contrasts this by delving into the role of design in teaching methods. At the same time, 'UTAUT Model and Technology Acceptance' examines the theoretical aspects of technology acceptance in education. The final themes, 'Knowledge Management and Learning Programs' and 'Reading Processes and Political Contexts in Education,' explore the strategic aspects of knowledge management in educational environments and the interplay between political narratives and educational content, respectively.

These topics are labelled according to the most frequent terms mentioned in Table 1, with the labels derived from a combination of the most weighted terms within each topic. These terms serve as indicators of the central themes being addressed in the collection of texts, thus guiding readers to grasp the core ideas quickly. Upon careful analysis, we observed that some topics bear similarities to others. We employed an intertopic distance map to delve deeper, as shown in Figure 4(a). This map, generated using dimensionality reduction techniques such as UMAP, visualises the distances between different topics identified in the dataset. It represents high-dimensional data in two or three dimensions, aiming to preserve relative distances as accurately as possible. This visualisation aids in distinguishing between topics and demonstrates the distinctness of each topic from the others. Well-separated clusters in the map suggest distinct topics, while areas where clusters overlap indicate similarity. Additionally, the map reveals relationships between topics, allowing us to infer potential connections or common themes. Topics in proximity may share specific keywords or relate to similar subject matter. This is a diagnostic tool to evaluate the quality of topic

modelling, where a good topic model is characterised by well-defined, non-overlapping clusters corresponding to coherent and distinct topics.

Figure 4(a) indicates some topics' similarities to others. We employed additional visualisations to validate these similarities further and understand the generated topics' pattern. We visualised this using a hierarchical clustering map in Figure 4(b) and a similarity matrix heatmap in Figure 5. The hierarchical clustering map, or dendrogram, visually groups similar topics together in a tree-like structure. Each branch in this structure represents a possible grouping, with branch lengths indicating the level of similarity between topics. Similar topics cluster under the same branch and are positioned closer, aiding in determining a cutoff level for merging related topics and thereby simplifying the topic model. In contrast, the similarity matrix heatmap visualises pairwise topic similarities in a matrix format, where warmer colours represent higher similarity and cooler colours have lower similarity. Each matrix cell shows the similarity score between two topics, with the diagonal indicating maximum similarity, as a topic is always 100% similar to itself. This heatmap is instrumental in quickly identifying topics that are similar or distinct. It also helps verify the coherence of the topics generated by the model, where coherent topics should exhibit higher similarity scores with themselves and lower scores with unrelated topics.

Therefore, based on these visualisations, we opted for a similarity threshold value of 0.83 to consolidate the topics. This thresholding led to four thematic categories, each characterised by a thematic similarity of 0.83 or higher. The details of these thematic groups are discussed in the following section, highlighting their significance and interrelations.

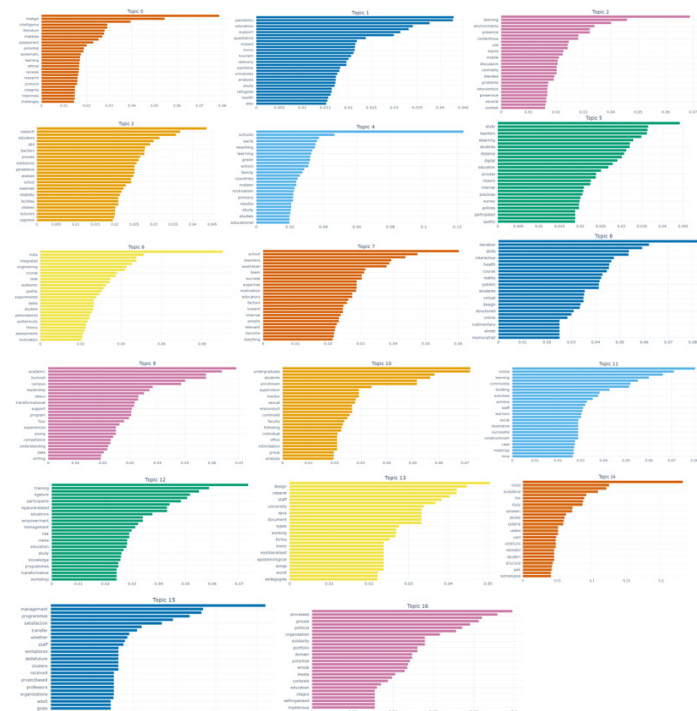


Figure 3. Topic word scores.

Table 1: Topics and their labels based on the most frequent terms.

Generated Topics	Created Labels	Most frequent terms
Topic 0	AI in Education	AI, ChatGPT, education, intelligence, artificial, literature, higher, chatbot, content, assessment, paper, potential, academic, systematic, generative, learning, human, ethical, benefits, reviews, using, research, tools, protocol, teachers, integrity, assessments, responses, impact, challenges
Topic 1	COVID-19 and Education	COVID, pandemic, online, education, students, support, higher, qualitative, research, impact, provided, NVivo, student, tourism, hybrid, delivery, wellbeing, part-time, model, university, future, analysis, impacted, study, graduate, refugees, data, health, identified, compassion
Topic 2	Online Learning Environments	online, learning, social, environments, simulation, presence, students, contentious, classes, use, interconnectivity, topics, using, mobile, distance, discussion, challenged, centrality, network, blended, levels, problems, various, intervention, books, preservice, include, several, multiple, context
Topic 3	Educational Feedback and Outcomes	feedback, research, instructional, educators, outcomes, students, teachers, academics, process, books, institutions, psychomotor, persistence, gender, analysis, factor, school, affective, materials, study, reliability, publishing, facilities, questionnaire, children, picture, lecturers, researchers, cognitive
Topic 4	Mathematics Education in Sub-Saharan Africa	mathematics, schools, Sub-Saharan, early, learners, teaching, ppd, Africa, learning, children, grade, teachers, school, family, quality, countries, Tanzania, Malawi, interventions, motivation, challenge, primary, curriculum, results, low, study, performance, studies, approaches, educational
Topic 5	Digital Learning Platforms and Practices	teaching, study, countries, teachers, school, e-learning, platforms, students, learners, distance, pd, digital, refugee, education, refugees, process, closures, closure, impact, internet, time, practices, used, survey, six, policies, access, participated, gap, quality
Topic 6	Integrated Programming and Engineering Equation	writing, India, students, integrated, programming, engineering, approach, course, thinking, task, skills, authentic, accounting, quality, engagement, experimental, developing, tasks, education, student, study, administered, findings, authenticity, problem-solving, theory, traditional, assessments, classroom, motivation
Topic 7	Teacher Professional Development and Learning	learning, school, routines, teachers, us, Australian, students, team, professional, success, motivating, expertise, teacher-led, motivation, experiences, educators, peers, factors, academic, toward, achievement, internal, teacher, people, development, relevant, study, become, knowledge, teaching
Topic 8	Clinical Skills and Healthcare Training	clinical, iteration, assessment, skills, game, interactive, nursing, health, application, course, user, reality, mixed, patient, experience, students, OSCEs, virtual, Singapore, design, serious, structured, resulted, online, realism, rudimentary, self-confidence, street, persons, memory trail
Topic 9	Academic Connectedness and Resilience	connectedness, academic, level, burnout, resilience, campus, PhD, leadership, students, EBSCO, findings, transformational, study, support, student, programme, graduate, empathy, bass, experiences, self-efficacy, young, individuals, competence, provide, understanding, kindness, counsel, within, humaneness
Topic 10	Research in Undergraduate Education	research, undergraduate, hdr, students, concurrent, enrolment, campus, supervision, matriculation, mentor, researchers, sexual, stem, misconduct, respondents, continued, university, faculty, opportunity, following, training, individual, member, office, eras, intimidation, Embry-Riddle, group, factors, analysis
Topic 11	Collaborative Online Teaching and Learning	teaching, online, collaborative, learning, collaboration, community, Jamboard, building, applied, activities, model, achieve, strong, staff, presence, learners, providing, social, deep, illustrative, scaffolded, successful, collaborativism, constructivism, due, case, support, meetings, contemporary, time
Topic 12	Healthcare Simulation and Training	healthcare, training, simulation-based, ligature, within, participants, health, ligature related, professionals, situations, ipe, empowerment, theme, management, care, risk, demonstrated, views, effectively, education, nursing, study, experiences, knowledge, themes, programmes, firsthand, transformative, transformed, workshop
Topic 13	Graphic Design and Modern Pedagogical Approaches in Higher Education	graphic, design, sbp, cabaret, modern, staff, pedagogic, university, kabarett, specs, thrs, document, ite, types, pedagogy, working, article, forms, critique, ironic, thr, neoliberalised, sociality, epistemological, regimes, songs, charged, world, performative, pedagogies
Topic 14	UTAUT Model and Technology Acceptance	utaut, model, use, acceptance, technology, like, educational, study, new, osmasem, theory, studies, empirical, systems, predicting, unified, intentional, used, initial, constructs, robust, extended, proven, equation, done, structural, framework, past, modelling, technological
Topic 15	Knowledge Management and Learning Programs	knowledge, management, motivation, programmes, project, satisfaction, design, transfer, learn, whether, learning, staff, academic, workplaces, Paulo, skillsfuture, elucidate, clusters, são, received, real life, project-based, intent, professors, distinct, organisations, three, adult, orientation, goals
Topic 16	Reading Processes and Contexts in Education	reading, processes, social, groups, neoliberalist, political, organisation, movements, solidarity, Aotearoa New Zealand, portfolio, critiques, domain, reforms, potential, particular, article, tertiary, media, forms, contexts, examine, education, tap, otagos, heart, self organised, suggestion, mysterious

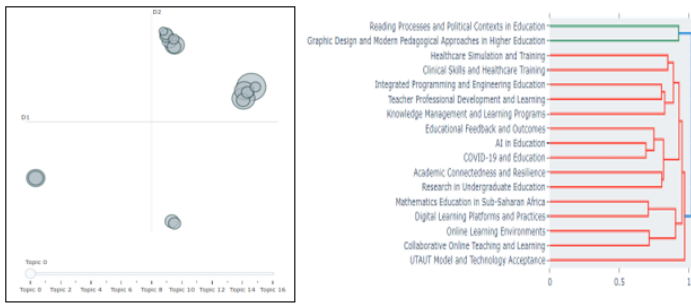


Figure 4. (a) Intertopic distance map (b) Hierarchical Clustering.

Compassionate Aspects in Education” group brings together themes like “COVID-19 and Education,” “Academic Connectedness and Resilience,” “Research in Undergraduate Education,” “Integrated Programming and Engineering Education,” “Knowledge Management and Learning Programs,” and “Reading Processes and Political Contexts in Education,” focusing on the educational impact of the pandemic.

These clusters categorise a broad range of topics into more concise and focused groups, each distinguished by its unique thematic focus. This clustering simplifies understanding of diverse topics and highlights distinct areas within the broader educational context.

## Discussion

Using topic modelling, we closely examined 17 topics that detail the trends of publications in the Journal of Applied Learning and Teaching. Subsequently, we extracted four thematic groups to explore the broader interests of researchers, particularly addressing various challenges in learning and teaching.

The first thematic group focuses on Technology and Digital Learning in Education, highlighting the growing significance of AI and technological innovations in education. As depicted in Figure 6, terms like ‘AI’, ‘ChatGPT’, and ‘artificial’ significantly emphasise the role of AI, especially the use of AI-driven tools such as chatbots in educational contexts. The frequent appearance of words like ‘online’, ‘e-learning’, and ‘blended’ indicates a shift towards digital and blended learning environments. This group also explores the ethical aspects of digital learning, including concerns about academic honesty and maintaining integrity in a digital academic landscape. Key terms such as ‘assessment’, ‘feedback’, and ‘engagement’ reflect a strong interest in evaluating the impact of these technologies, particularly regarding educational outcomes and student involvement. Figure 7 further illustrates a consistent increase in the importance of technological and digital tools in educational settings, aligning with the study (Rudolph et al., 2023). The group envisions a future where digital platforms are central to content delivery and student engagement. However, it also points to the unethical use of tools like ChatGPT in assessments, which breaches ethical standards. Solutions include constructive feedback, rigorous supervision, interactive methods, and teamwork activities to ensure adherence to ethical norms in the evolving educational landscape.

The second thematic group, ‘Healthcare and Clinical Training,’ emphasises the crucial role of practical, experiential learning in healthcare education. As shown in Figure 6, there is a distinct focus on healthcare education, particularly in nursing and clinical disciplines. Terms like ‘clinical’, ‘nursing’, ‘skills’, and ‘assessment’ are frequently mentioned, underscoring the importance of hands-on skill development and evaluation in this sector. Notably, using terms such as ‘virtual’, ‘interactive’, ‘reality’, and ‘game’ indicates the adoption of advanced, technology-driven training methods such as virtual reality and interactive simulations. This innovative approach aims

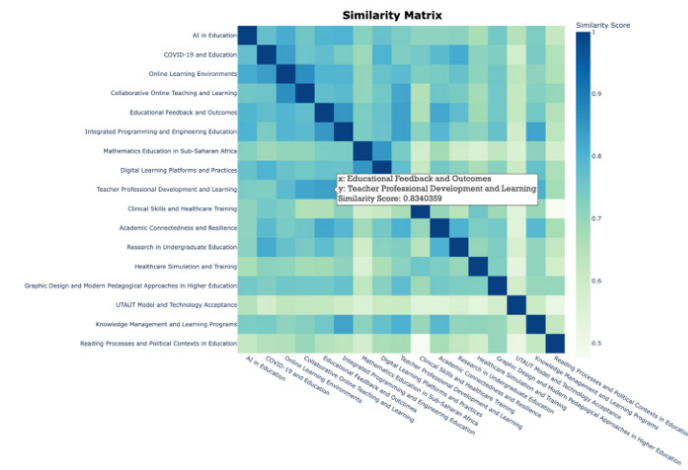


Figure 5. The heatmap presented a similarity matrix between the topics.

## Thematic groups

We employed a clustering method focused on thematic resemblance to form coherent clusters from the identified topics. This involved an in-depth examination of each topic, merging those with intersecting or related themes, following the methodology suggested by Glazkova (2021). We used a benchmark similarity score of 0.83 for grouping, ensuring high relevance and cohesion within each cluster. Subsequently, descriptive labels were carefully selected for each cluster, aiming to capture the essence of the combined topics succinctly.

The first thematic group, “Technology and Digital Learning in Education,” unites topics like “AI in Education,” “Online Learning Environments,” “Digital Learning Platforms and Practices,” and “UTAUT Model and Technology Acceptance.” This cluster reflects the increasing integration of technology in educational settings. The second group, titled “Healthcare and Clinical Training,” encompasses “Clinical Skills and Healthcare Training” along with “Healthcare Simulation and Training.” This cluster highlights the importance of hands-on training and simulation in healthcare education. The third group, “Educational Strategies and Outcomes,” includes a diverse array of topics such as “Educational Feedback and Outcomes,” “Teacher Professional Development and Learning,” “Mathematics Education in Sub-Saharan Africa,” “Collaborative Online Teaching and Learning,” and “Graphic Design and Modern Pedagogical Approaches in Higher Education,” reflecting various educational methods and outcomes. Lastly, the “Pandemic-Driven Social and

to enhance learning and increase healthcare practitioners' confidence. Furthermore, the presence of terms like 'gender', 'equality', and 'female' in this thematic group suggests attention to gender perspectives in healthcare training, potentially addressing issues of diversity and inclusiveness. Essentially, this group reflects a trend towards more engaging and immersive learning experiences through simulations and virtual reality, which is particularly vital in the dynamic and constantly evolving healthcare sector. The need for continuous training and skill enhancement became especially prominent during the COVID-19 pandemic. As depicted in Figure 7, the peak interest in this area occurred around 2021, coinciding with the heightened impact of the COVID-19 pandemic, indicating an intensified focus on healthcare training during this period.

The third thematic group, 'Educational Strategies and Outcomes,' delves into various educational tactics and their impacts. The terminology depicted in Figure 6 paints a broad picture of education and learning. Terms such as 'feedback', 'instructional', 'outcomes', and 'pedagogy' highlight a focus on teaching methods, student evaluation, and education results. Regular references to 'students', 'teachers', 'educators', and 'learners' emphasise the central figures in the educational process. Furthermore, words like 'research', 'study', 'analysis', and 'publishing' suggest a scholarly approach to education, potentially encompassing academic research and the field of educational publishing. The mention of topics like 'mathematics', 'cognitive', 'affective', and 'psychomotor' indicates an exploration of various learning domains and subject areas, showing a wide array of educational interests and approaches. This theme's distribution, as seen in Figure 7, indicates steady progress and possibly modest enhancements in educational strategies and outcomes. Overall, this group underscores the critical role of teacher development in elevating educational quality. The specific focus on mathematics education in Sub-Saharan Africa signals attention to region-specific educational challenges and tailored solutions, including innovative pedagogical methods and graphic design, pointing to a shift towards more dynamic and engaging teaching techniques.

Finally, the fourth thematic group, 'Pandemic-Driven Social and Compassion Aspects in Education,' sheds light on the societal and emotional dimensions influencing education during the pandemic (Figure 6). This group, marked by terms like 'COVID', 'pandemic', 'online', 'education', 'wellbeing', and 'hybrid', captures the diverse ways the COVID-19 pandemic reshaped educational practices. It includes the shift to online and hybrid learning models, emphasising the necessary adjustments in educational approaches during this period. The focus on student welfare, mental health, and resilience reflects students' psychological challenges. Additionally, this group highlights the pandemic's significant impact on higher education, with terms like 'university', 'graduate', and 'PhD' indicating changes in teaching, learning, and research dynamics. Including research-oriented terms implies investigating the pandemic's effects through qualitative studies and data analysis. Moreover, this group underscores the importance of maintaining community bonds and academic connectedness despite physical distancing. Keywords like 'skillsfuture' and 'mentoring'

emphasise skill enhancement and career preparation, adapting to changing job markets. The use of terms like 'connectedness', 'compassion', 'empathy', 'community', 'motivation', and 'kindness' highlights the importance of social and emotional support during these challenging times. As shown in Figure 7, interest in this thematic area showed a sharp increase up to 2022, followed by a decline, suggesting a peak in concern for the pandemic's social and compassionate aspects in education, which appears to have decreased by 2023. Overall, this thematic category offers a holistic view of the educational sector's adaptation to the pandemic, emphasising digital transformation, emotional support mechanisms, and preparation for future challenges.

Each thematic group provides a lens through which to understand current trends, challenges, and future directions in education and training. They collectively highlight the dynamic nature of education, the growing influence of technology, and the importance of contextual and culturally sensitive approaches.



Figure 6. Word clouds of thematic groups.

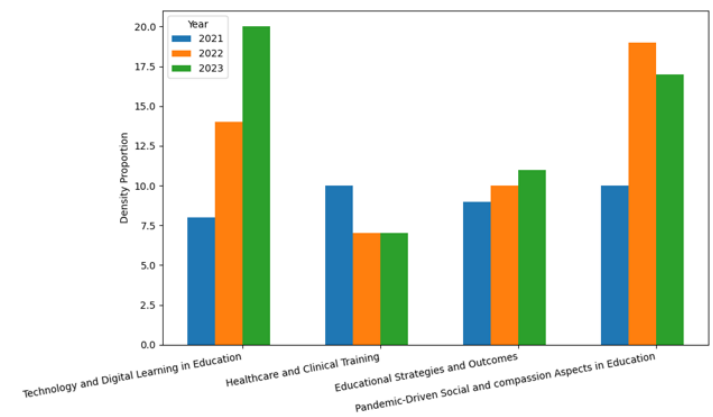


Figure 7. Year-wise densities distribution of thematic groups.

## Conclusion

In conclusion, applying BERTopic modelling to the Journal of Applied Learning and Teaching has provided critical insights into the evolving landscape of higher education in the post-COVID era. Through a detailed analysis of academic texts from January 2021 to December 2023, we identified key research themes and trends within higher education. We meticulously organised 17 predominant topics into four broad thematic groups: Technology and Digital Learning in Education, Healthcare and Clinical Training, Educational



Strategies and Outcomes, and Pandemic-Driven Social and Compassionate Aspects in Education. These groups highlight the journal's interdisciplinary approach and offer a comprehensive understanding of the shifts in educational practices and research focus areas. Our annual examination of these themes, illustrated through various data visualisations, underscores the dynamic challenges and opportunities identified by researchers. The study emphasises the growing importance of technology and digital learning, the intricacies of healthcare training, the efficacy of educational strategies, and the social and emotional impacts of the pandemic, all of which are pivotal in shaping the future trajectory of higher education. This rigorous analysis directs attention to crucial areas for further research and development, which hold the potential to refine applied learning and teaching methodologies significantly. As the educational sector continues to evolve in the aftermath and beyond the pandemic, this study is a strategic foundation for educators and policymakers to enhance resilience and adaptability. It provides invaluable guidance for addressing the ongoing global transformations in education, ensuring that the sector remains responsive and forward-thinking in these changing times.

## References

Ai, B. (2017). Constructing an academic identity in Australia: An autoethnographic narrative. *Higher Education Research & Development, 36*(6), 1095-1107. <https://doi.org/10.1080/07294360.2017.1303459>

Alasadi, S. A., & Bhaya, W. S. (2017). Review of data preprocessing techniques in data mining. *Journal of Engineering and Applied Sciences, 12*(16), 4102-4107. <http://dx.doi.org/10.3923/jeasci.2017.4102.4107>

Amado, A., Cortez, P., Rita, P., & Moro, S. (2018). Research trends on big data in marketing: A text mining and topic modeling-based literature analysis. *European Research on Management and Business Economics, 24*(1), 1-7. <https://doi.org/10.1016/j.iedeen.2017.06.002>

Bala, I. (2023). Natural language processing in medical science and healthcare. *Medicon Medical Sciences, 4*, 01-02. <https://doi.org/10.55162/MCMS.04.088>

Bala, I., Kelly, T.-L., Lim, R., Gillam, M. H., & Mitchell, L. (2023). An effective approach for multiclass classification of adverse events using machine learning. *Journal of Computational and Cognitive Engineering*. <https://doi.org/10.47852/bonviewJCCE32021924>

Bala, I., Kelly, T.-L., Stanford, T., Gillam, M. H., & Mitchell, L. (2024). Machine learning-based analysis of adverse events in mesh implant surgery reports. *Social Network Analysis and Mining, 14*(1), 63. <https://doi.org/10.1007/s13278-024-01229-6>

Cao, Q., Cheng, X., & Liao, S. (2023). A comparison study of topic modeling-based literature analysis by using full texts and abstracts of scientific articles: A case of COVID-19 research. *Library Hi Tech, 41*(2), 543-569. <https://doi.org/10.1108/LHT-03-2022-0144>

[org/10.1108/LHT-03-2022-0144](https://doi.org/10.1108/LHT-03-2022-0144)

Chen, X., Zou, D., Cheng, G., & Xie, H. (2020). Detecting latent topics and trends in educational technologies over four decades using structural topic modeling: A retrospective of all volumes of computers & education. *Computers & Education, 151*, 103855. <https://doi.org/10.1016/j.compedu.2020.103855>

Choi, J., & Lee, B. (2023). Quantitative topic analysis of materials science literature using natural language processing. *ACS Applied Materials & Interfaces, 16*(2), 1957-1968. <https://doi.org/10.1021/acsami.3c12301>

Escueta, M., Quan, V., Nickow, A. J., & Oreopoulos, P. (2017). Education technology: An evidence-based review. *NBER Working Paper Series*. <https://doi.org/10.3386/w23744>

Glazkova, A. (2021). Identifying topics of scientific articles with BERT-based approaches and topic modeling. In M. Gupta & G. Ramakrishnan (Eds.), *Trends and applications in knowledge discovery and data mining*. Springer International Publishing. [https://doi.org/10.1007/978-3-030-75015-2\\_10](https://doi.org/10.1007/978-3-030-75015-2_10)

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. arXiv preprint arXiv:2203.05794. <https://doi.org/10.48550/arXiv.2203.05794>

Gurcan, F., & Cagiltay, N. E. (2023). Research trends on distance learning: A text mining-based literature review from 2008 to 2018. *Interactive Learning Environments, 31*(2), 1007-1028. <https://doi.org/10.1080/10494820.2020.1815795>

Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., & Mathur, I. (2016). *Natural language processing: Python and NLTK*. Packt Publishing Ltd.

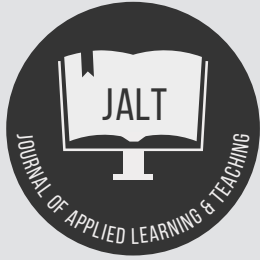
Hu, X., & Zhang, R. (2022, June). Text classification based on machine learning. In *2022 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* (pp. 911-916). IEEE. <https://doi.org/10.1109/ICAICA54878.2022.9844556>

Hussain, S., Ayoub, M., Jilani, G., Yu, Y., Khan, A., Wahid, J. A., ... & Weiyan, H. (2022). Aspect2Labels: A novelistic decision support system for higher educational institutions by using multi-layer topic modelling approach. *Expert Systems with Applications, 209*, 118119. <https://doi.org/10.1016/j.eswa.2022.118119>

Huston, C. L., Phillips, B., Jeffries, P., Toderro, C., Rich, J., Knecht, P., ... & Lewis, M. P. (2018, January). The academic-practice gap: Strategies for an enduring problem. *Nursing Forum, 53*(1), 27-34. <https://doi.org/10.1111/nuf.12216>

Jiang, S., Prasad, A., Kan, M. Y., & Sugiyama, K. (2018, August). Identifying emergent research trends by key authors and phrases. In *Proceedings of the 27th international conference on computational linguistics* (pp. 259-269). <https://aclanthology.org/C18-1022>

- Lemay, D. J., Baek, C., & Doleck, T. (2021). Comparison of learning analytics and educational data mining: A topic modeling approach. *Computers and Education: Artificial Intelligence*, 2, 100016. <https://doi.org/10.1016/j.caeai.2021.100016>
- Maphosa V., & Maphosa M. (2023). Artificial intelligence in higher education: A bibliometric analysis and topic modeling approach. *Applied Artificial Intelligence*, 37(1), <https://doi.org/10.1080/08839514.2023.2261730>
- McInnes, L., Healy, J., & Astels, S. (2017). HDBSCAN: Hierarchical density based clustering. *The Journal of Open Source Software*, 2(11), 205. <https://doi.org/10.21105/joss.00205>
- McInnes, L., Healy, J., & Melville, J. (2018). *Umap: Uniform manifold approximation and projection for dimension reduction*. arXiv preprint arXiv:1802.03426. <https://doi.org/10.48550/arXiv.1802.03426>
- Nylander, E., Fejes A., & Milana M. (2022). Exploring the themes of the territory: A topic modelling approach to 40 years of publications in International Journal of Lifelong Education (1982–2021). *International Journal of Lifelong Education*, 41(1), 27-44. <https://doi.org/10.1080/02601370.2021.2015636>
- Pandur, M. B., Dobša, J., & Kronegger, L. (2020, January). *Topic modelling in social sciences: Case study of web of science* [Conference Paper]. Central European Conference on Information and Intelligent Systems, Varazdin. <https://www.proquest.com/openview/f5b35b8664922bcf629d4a57a6344a45/1?pq-origsite=gscholar&cbl=1986354>
- Ranjbar-Sahraei, B., & Negenborn, R. R. (2017). *Research positioning & trend identification: A data-analytics toolbox*. TU Delft. <http://resolver.tudelft.nl/uuid:e736746b-7cba-4239-ac87-49634ccb43e4>
- Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence embeddings using Siamese BERT-networks*. arXiv preprint arXiv:1908.10084. <https://doi.org/10.48550/arXiv.1908.10084>
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Wang, Z., Chen, J., Chen, J., & Chen, H. (2023). Identifying interdisciplinary topics and their evolution based on BERTopic. *Scientometrics*, 1-26. <https://doi.org/10.1007/s11192-023-04776-5>
- Yunita, S. (2018). Theme and thematic progression in students' recount texts. *Indonesian Journal of Applied Linguistics*, 7(3), 524-530. <https://doi.org/10.17509/ijal.v7i3.9797>



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Is ChatGPT an opportunity or a threat? Preventive strategies employed by academics related to a GenAI-based LLM at a faculty of education

---

Micheal M Van Wyk<sup>A</sup>

A

*University of South Africa*

---

### Keywords

---

Academic dishonesty;  
artificial intelligence;  
constructivist-interpretative perspective;  
exploratory qualitative design;  
GenAI tools.

### Abstract

---

Within the past decade, enormous strides have been made related to the disruptive effect of AI in education, which has grown exponentially. Recent developments in GenAI conversational models have highlighted the need to investigate this phenomenon in different contexts. This prompted me to investigate academics' views on ChatGPT as a GenAI-based conversation tool at a faculty of education. The conversation theory is foregrounding this research. An exploratory qualitative design study foregrounded the constructivist-interpretative perspective and a sample of eight participants was purposively selected. Semi-structured interviews were generated by Microsoft Teams (transcribed), recordings were downloaded, and themes were identified as guided by the thematic analysis process. Participants echoed sentiments of the usefulness of generative AI tools in promoting or advancing teaching and learning experiences. An awareness of the ethical considerations in using generative AI tools is important before adopting chatbots. To prevent the unethical behaviour of students, it is necessary to create and adopt measures to prevent academic dishonesty. Further research is needed to build on recent gains in academic awareness of GenAI tools for teaching and learning.

---

### Correspondence

[vwykmm@unisa.ac.za](mailto:vwykmm@unisa.ac.za)<sup>A</sup>

---

### Article Info

Received 11 November 2023

Received in revised form 9 January 2024

Accepted 13 February 2024

Available online 19 February 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.15>

## Introduction

Artificial intelligence (AI) is not a recent concept in research. Consider, for instance, Isaac Asimov's (1955) famous robotic invention based on the "Three Laws of Robotics", and the first science fiction book, "I Robot". Furthermore, Allan Turing (1950) extended his scholarly work on AI research, which prompted UNESCO to recommend the Ethics of Artificial Intelligence (AI) in the application of machine learning (robots) to prevent abuse, fraud, and security risks. Based on the principles of ethics of AI, different AI tools were developed to minimise academic integrity risks. Those early years of exploring AI in higher education opened many opportunities and challenges to the sector. The rapid development of AI research has a profound impact on higher education. In the past decade, reports revealed that leveraging on the advances in AI-powered solutions has carried enormous benefits to accelerate the UNESCO Sustainable Development Goals and the African Agenda 2063 Strategy (Goralski & Tan, 2023; Kamalov et al., 2023; Abulibdeh et al., 2024).

The turning point in November 2022 was the launch of the large language model, ChatGPT 3.5 (Browne, 2023), a sophisticated conversational tool based on artificial intelligence (AI) by OpenAI, which created an unprecedented movement globally. The characteristics of the infamous updated version, GPT-3.5, which was developed by tech company OpenAI, created a "hype" for the generative AI conversational (GenAI) tool, which responds to user text prompts that are indistinguishable from human actions. Several other GenAI prompting engineering software emerged, such as Claude2 and Llama2, which sparked the so-called "California Dollar Rush" in Silicon Valley (Rudolph et al., 2023; Griffith & Metz, 2023). However, Nemorin et al. (2023) raise a concern that "*many of the claims of the revolutionary potential of AI in education are based on conjecture, speculation, and optimism*" (p. 39). Scholars claimed that amidst the tech war amongst USA tech companies and the hype around different types of chatbots, ChatGPT grew faster than any other chatbot as well as social media platforms like Twitter, WhatsApp, or Instagram (Wodecki, 2023; Yang, 2022). Leveraging the potential of GenAI in education, Holmes, Bialik and Fadel (2019) argue that current GenAI initiatives and other chatbots as large language models (LLMs) have impacted education significantly and have grown exponentially. Given the latter, several scholarly works report on the impact of GenAI on tuition, research, assessment, ethics, deep learning and professional development at higher education institutions (Baidoo-Anu & Owusu Ansah, 2023; Smolansky et al., 2023; Vargas-Murillo et al., 2023; Dwivedi et al., 2023; Wang & Zhang, 2023). In spite of claims of the revolutionary potential of AI in education, concerns are raised about ethical issues and academic integrity (Pisica et al., 2023; Al Matari et al., 2023; Baidoo-Anu & Owusu Ansah, 2023). Some argue that this phenomenon has yet to explore whether GenAI tools have the potential to track learning outcomes across contexts and competencies. At the time of exploring academic staff's awareness of GenAI conversational LLM in teaching and learning (Popenici & Kerr, 2017), a few studies have been conducted at higher education institutions (HEIs) (Smolansky et al., 2023; Peres et al., 2023) but to a lesser

extent in the African context (Mhlanga, 2023; Baidoo-Anu & Owusu Ansah, 2023; Adarkwah et al., 2023; Van Wyk et al., 2023).

Recent developments in GenAI initiatives and other chatbots as large language models (LLMs) have highlighted the need to explore ChatGPT as a GenAI phenomenon in higher education in general, but specifically for a teaching and learning context. I was prompted to investigate academics' experiences of using ChatGPT as a GenAI-based conversation chatbot in teaching and learning at a faculty of education. Based on this exploratory inquiry, the following research questions (RQ) were formulated:

- RQ1: What are the views of academics about GenAI-based conversational chatbots in teaching and learning at a faculty of education?
- RQ2: What measures or strategies do the academics apply to prevent academic cheating of using GenAI-based chatbots by students in teaching and learning at a faculty of education?

## Literature review

### Conversation theory underpinning the study

In 1966, Gordon Pask delivered his seminal text on the cybernetic model based on learning and competencies at the Ohio Bionics Symposium. This seminal text supported his co-authored publication with Pask and Scott (1972), as an extended version of his cybernetic model based on man-machine learning strategies and competence. This article is underpinned by the Conversation Theory (CT) developed by Gordon Pask (1976) and applied in the context of the hype created by the launch of ChatGPT by OpenAI. The notion is that CT is based on the three constructs of cognition, conversation, and learning. The CT functions on concept-forming and concept-sharing between conversational participants.

I concur that ChatGPT3.5 is a conversational chatbot [cloud-based robot], programmed to understand, interpret, and start conversations with humans [sic...]. It depends on the type of prompting or interactive conversations with the robot-in-the-cloud [chatbot]. Therefore, ChatGPT is a conversational LLM chatbot, which acts on prompting. So why prompt engineering? ChatGPT functions optimally on prompts to generate text. It is computerised as an LLM, based on processes to generate text for understanding and interpretation. Studies reported that academics and students are aware of the benefits and drawbacks of ChatGPT as a conversational LLM chatbot (Megahed et al., 2023; Rudolph et al., 2023). Moreover, research showed that ChatGPT is a conversational robot based on natural language processing (NLP) which engages as a robot with users in a human-like conversation (Adarkwah et al., 2023; Kamalov et al., 2023). Therefore, ChatGPT is a conversational LLM chatbot that supports, creates, and facilitates meaningful interactions and conversations with humans (Wodecki, 2023).

## The impact of GenAI-based research on the future of education

Based on the ground-breaking AI work by torchbearers Asimov and Turing, their research has led to tremendous strides in this area and has profound impact on the field of education. This has led to the use of AI tools to enhance teaching and learning experiences. The launch of the first LLM model created hype around ChatGPT as a GenAI-based conversational chatbot. In a newspaper article by Wayne Hu (2023) entitled *"ChatGPT sparks AI 'gold rush' in Silicon Valley"*, the author made a profound claim that *"this wave of AI could be bigger than mobile or the cloud, and more on the scale of something like the Industrial Revolution that changed the course of human history"* (p. 1). Similarly, scholars coined the phrase "California Dollar Rush" to describe (Rudolph et al., 2023; Griffith & Metz, 2023) the frenzy among tech companies to have a stake in the GenAI tech war. This has increased the complexity of prompt engineering paradoxical sentiments and has impacted the future of higher education and business (Oxford Analytica, 2023).

So, this leaves us with an unanswered question, namely whether ChatGPT and similar chatbots should be tagged as either a "friend or a foe" (Pisica, 2023; Matari et al., 2023). I believe that GenAI is the future and faculties of education must ride the wave of GenAI. Moreover, universities can ride the "hype" of popularity created by tech companies. Universities can leverage the amazing capabilities of GenAI chatbots (Bozkurt et al., 2023) and take advantage of creating innovative teaching and learning opportunities without compromising values such as ethics, diversity, equality, social justice, and quality education (Holmes et al., 2019). Research reported that GenAI-based LLM chatbots, for example, ChatGPT, were useful chatbots that generated specific content knowledge and supported specific writing tasks for both lecturers and students (Kaplan-Rakowski et al., 2023; Megahed et al., 2023; Michel-Villarreal et al., 2023). Studies exploring the use of GenAI in education revealed tremendous advances and growth in the teacher education context (Holmes et al., 2019; Wodecki, 2023; Su & Yang, 2023). Studies showed that, for example, ChatGPT and Claude2 not only acted on prompts in a "humanly" manner but could also identify student learning gaps (Lim et al., 2023). On the other hand, ChatGPT is a large language technological robot that uses an instructional approach referred to as prompt engineering (Rose et al., 2023).

In sum, the conversation theory foregrounds this article, namely that ChatGPT is an exciting and value-added conversational chatbot that academics can use to advance teaching and learning spaces. The chatbot brought profound changes related to the future of education. Therefore, lecturers have opportunities to reimagine their praxis because of the increased popularity of chatbots. Lecturers need to rethink the purpose of using chatbots for teaching and learning.

## The popularity of ChatGPT as a conversational chatbot at a faculty education

Recently, GenAI in education, in particular ChatGPT, has created hype in teacher education (Yang, 2022; Mohamed, 2023; Antonenko & Abramowitz, 2023). Some of these studies were conducted in contact, blended learning, and open-distance learning contexts (Chan & Hu, 2023; Baidoo-Anu & Owusu Ansah, 2023; Adarkwah et al., 2023; Van Wyk et al., 2023). As reported, GenAI-conversational chatbots showed exponential promise for teacher education. Furthermore, studies conducted in these contexts reported that these chatbots offer pedagogical possibilities for student learning and teaching opportunities (Baek & Kim, 2023; Chan & Hu, 2023; Adarkwah et al., 2023). It could be deduced that the GenAI phenomenon has brought a pedagogical shift in teacher education. For example, lecturers viewed the efficacy of ChatGPT in teaching English Foreign Language students (Mohamed, 2023). Research reported the advantages of using GenAI tools in fostering a student-centred approach (Huang et al., 2022; Antonenko & Abramowitz, 2023). Studies reported that GenAI tools support personalised learning experiences (Chan & Hu, 2023; Li & Wong, 2023; Lodge et al., 2023). It also fulfils students' cognitive needs (Baek & Kim, 2023; Yang, 2022). In the context of this study, academics were exposed to ChatGPT, and they frequently used it for teaching and learning in the faculty of education. To increase the popularity of chatbots for student learning, research reported that teachers implemented innovative approaches to assessing student work in different subjects (Van Wyk et al., 2023; Antonenko & Abramowitz, 2023; Smolansky et al., 2023; Peres et al., 2023).

## Ethical dilemmas and academic integrity using GenAI technologies in teaching and learning

Although GenAI tools are available for academics and students, the threat of academic dishonesty, cheating, plagiarism, and copyright infringement has ethical implications for education (Jarrah et al., 2023; Gao et al., 2023; Baek & Kim, 2023; Mhlanga, 2023; Vargas-Murillo et al., 2023; Eke, 2023; Peres et al., 2023). As reported in the latter studies, universities need to revise specific policies related to tuition, research, and assessment practices to prevent cheating and preserve academic integrity at all costs. An empirical study by Cooper (2023) reported key concerns regarding ethical considerations related to the use of copyright infringements by science teachers and students, which could compromise academic integrity. To curb this challenge, academics must be capacitated to use GenAI tools to raise awareness and set an example for students to follow. Studies reported possible remedies to be used by academics on the use of GenAI tools in their praxis (Ali, 2023; Huallpa, 2023; Dwivedi et al., 2023; Cotton et al., 2023; Kelly et al., 2023). Research reported that students use ChatGPT and similar chatbots to cheat academically (Jarrah et al., 2023). In view of the latter challenge, a concern is raised that university policies are "silent" on academic cheating at some faculties of education. Students exploit this policy gap and thus cannot be penalised or sanctioned for plagiarism or copyright infringements for using ChatGPT technologies and similar LLMs in academic writing (Gao et

al., 2023). Therefore, to address the “chatbot in the cloud”, I assumed that universities advocate the seriousness of cheating as well as empower students with competencies to use LLM technologies ethically. Therefore, institutions of higher education have a moral duty to popularise preventive strategies to remedy cheating, plagiarism and copyright infringements and advance academic integrity (Mijwil et al., 2023).

### Strategies to prevent academic cheating with GenAI tools

Scholars argued that to protect the image and integrity of the institutions of higher learning, specific measures must be developed to prevent cheating (Kumar & Mindzak, 2024; McDonald et al., 2024). Chan (2023) is of the view that ethical considerations must be applied and, if needed, policies related to integrity, copyright infringements and praxis of academics should be revised. Given the latter, institutions of higher learning can implement strategies to prevent academic cheating. Literature provides measures or strategies that can be considered to prevent academic cheating of generative AI tools in teaching and learning. Studies by Oravec (2023) and Firat (2023) reported specific cheating-detection strategies and GenAI-empowered skills that could be used to support students in learning to deal with academic dishonesty or plagiarism. Lecturers must advocate and create an awareness of why and how generative AI can be viewed as a learning opportunity to increase competence in GenAI tools. As a measure to prevent cheating or academic dishonesty, students could be given case studies (Sallam, 2023), research- and problem-based projects (Firaina & Sulisworo, 2023) to present the results and have them write and report as a group. Strategies such as group discussions (Castillo et al., 2023), critical conversation forums and online panels are also proposed. These strategies will help to raise awareness of and enhance critical conversations about GenAI technologies to protect academic integrity and freedom and instil a sense of integrity in students’ work.

### Methodology

Before the study began, an invite was sent to identified participants. During the year, academics were exposed to several webinars on generative AI tools and conferences. Based on this exposure, the participants agreed and signed a consent application to take part in the study. I foregrounded this exploratory qualitative design study from the constructivist-interpretative perspective and sought the views of academics who explored the generative AI-based conversational LLM models. After approval of ethical clearance (Ref EFEC 5-08/2023), an invite was posted to all academics at the faculty of teacher education, an institution of higher learning. The purpose of the exploratory study on ChatGPT was stated. To be selected as participants in the study, specific criteria were stated, namely awareness of GenAI tools *{awareness and early adaptors}* and applying ChatGPT as a GenAI tool in teaching and learning *{adaptability and accessibility}*. For the purpose of selecting the sample for this study, an invite was sent to 42 academic staff in teacher education (faculty of education) at an institution

of higher education. The invite requested academic staff to indicate exposure to GenAI technologies through attending conferences, webinars, in-service training or currently used GenAI technologies for teaching and learning. Based on pre-data collection, a quota sample of 9 participants (9/42 = 21.4%) were selected who “met” the criteria to be selected for the study. All participants signed a consent application. Based on information obtained through the pre-interview session (criteria for selection), the faculty of education hosted several webinars, panel discussions and in-service training in GenAI technologies. The purposive quota sampling comprised both males (33.7%) and females (66.7%) academics at a selective institution of higher learning (see Table 1).

Table 1. Biographical information of participants (n=9).

Participants	Code for participants	Gender	Lecturing experience (in Years)
Lecturer 1	L1	Female	8
Lecturer 2	L2	Female	6
Senior Lecturer 1	SL1	Female	9
Senior Lecturer 2	SL2	Female	7
Associate Professor 1	AP1	Female	14
Associate Professor 2	AP2	Male	19
Full Professor 1	FP1	Male	15
Full Professor 2	FP2	Male	23
Manager: Teaching and Learning	MTL	Female	19

The quota sample received dates and times with a link scheduled on Microsoft Teams. All semi-structured interviews were generated by Microsoft Teams {transcribed}, and recordings were downloaded and secured as part of data management. To ensure the trustworthiness of the generated themes reported, the transcripts, Microsoft Teams recordings, and semi-structured interview protocol were sent to the quota sample to verify and validate, through member checking, the correctness of the data generated. An online link sent the recordings, transcripts, and verified extracts and questions posted by participants for them to verify their correctness. After a week, if participants agreed, an e-mail was sent to the study’s quota sample. The extracts were analysed manually, and several themes were identified as guided by the thematic analysis process (Creswell, 2012; Nowell et al., 2017).

### Findings

The results confirm that, to date, little research has been done using a systematic review based on the methodological framework of ChatGPT as an AI conversational tool. Academics were aware of the potential benefits and drawbacks of the usefulness and functionality of ChatGPT for teaching and learning. Further investigations are recommended to explore similar studies in teacher education.

**RQ1: What are the views of academics about GenAI-based conversational chatbots in teaching and learning at an institution of higher education?**

After data analysis, several themes emerged from the data:

- Using generative AI tools in promoting and/or advancing teaching and learning experiences of students.
- Generative AI conversational tools such as ChatGPT are likely changing the face of higher education.
- Creating awareness and ethical considerations for the use of generative AI tools.

### Using generative AI tools in promoting and/or advancing teaching and learning experiences of students

The majority of participants (90%) in the study were positive about the usefulness of the generative AI apps, while 10% were concerned about cheating and academic dishonesty. Furthermore, 80% of them were registered and used the free version of ChatGPT-3.5 for teaching and learning. Participants viewed generative AI conversation modules as a prompt engineering tool that benefitted teaching and learning. One participant said she used specific case studies to promote problem-solving skills and academic writing opportunities. This participant said:

"I am using ChatGPT as a prompt engineering tool to generate specific case studies. I use these case studies to promote problem-solving skills and academic writing opportunities. For example, my master's degree students received a ChatGPT-generated case study, they evaluated the content, in-text referencing, language editing (grammatical errors), paraphrased the case study and presented each case study for the group to critique their presentations" (FP2).

Another participant found ChatGPT useful, capacitating postgraduate students with academic writing skills. She echoed:

"I found one specific AI tool, ChatGPT, very useful for helping my students with their academic writing. In my view, GenAI tools can be banned, embraced, ignored, designed around it or go back to traditional assessment practices. But gives it a chance to evaluate where it is useful and reliable for your context" (AP1).

### Generative AI conversational tools such as ChatGPT are likely changing the face of higher education

Participants believed that the hype about ChatGPT and other LLM chatbots is here to stay, but higher education must adjust their policies and guidelines related to assessment, plagiarism, academic dishonesty, and ethics. Participants agree that generative AI tools have disrupted the sector and are likely to change the face of education. L1 opined: *"When ChatGPT was launched last year, a hype was created, and disrupted the higher education sector. Higher education needs to adjust policies such as ethics, exams, and assessment"*.

Most participants had read an article or attended a webinar or conference on ChatGPT. AP1 said: *"Since the day I heard about ChatGPT, everybody, including my students, has been*

*experimenting with this tool."* However, SL2 is of the view that tech companies have invested to gain an advantage in AI-generative apps for profit-making. She said: *"Last week, I read an article about the so-called 'California Dollar Rush'. Tech companies invested millions of dollars in generative LLM chatbots. One tech company, OpenAI, invested millions of dollars. A tech war erupted among US companies to see who claimed the most dollars"* (SL2).

### Creating awareness of the usage of GenAI conversational tools and the ethical implications of using GenAI tools

Participants were cognisant of how fast the generative AI tools emerged since the first chatbot was launched. They said academics must be mindful of the advantages and drawbacks of AI conversational models before jumping to use them in teaching and learning. FP1 is aware of the tools available: *"Several others have been developed, Claude 2 and Llama2. There is a war among tech companies to get a bigger slice of the chatbot pie. But what about the ethical implications. Are we addressing the real issue of cheating"*. Many mentioned that the "elephant in the cloud" is about ethics. In addition, several participants raised concerns about ethics and how to detect cheating or plagiarism.

"ChatGPT has sparked heated debates around ethical issues like academic dishonesty and cheating by both academics and students. But there are also generative AI-detection tools to detect academic dishonesty and cheating. I used Turnitin as well GPTZero for tracking generative content in assignments. ChatGPT can easily generate any text or an assignment or even a research proposal" (SL1).

According to this participant, another major issue to be addressed is revising the existing policies related to research, assessment, and ethics as a matter of urgency. *"As an institution, stakeholders were informed of the revision of policies related to tracking of an assignment or text was written by a generative AI software. An awareness was created and policies such as teaching and learning, ethics, assessment, academic integrity, and research were approved"* (L3).

### RQ2: What measures or strategies do the academics apply to prevent academic cheating by students using GenAI-based chatbots in teaching and learning?

Different themes emerged from the data analysis process, discussions, and specific extracts from participants:

- Combat academic fraud, cheating and dishonesty through authentic assessment/alternative assessment tasks.
- Empower students with digital literacy skills in using GenAI detection tools as a preventive measure to discourage cheating.
- Advocate and inculcate principles of integrity, morality, and ethical responsibility in using GenAI-chatbot strategies in academic writing.

- Design authentic assessment tasks that cannot easily be generated by GenAI chatbots but are applied to real-life contexts for meaningful learning.

### **Combat academic fraud, cheating and dishonesty through authentic assessment/alternative assessment tasks**

Most of the participants (90%) agreed that it is necessary to create and adopt measures to prevent academic fraud, cheating and dishonesty. They view the purpose of using chatbots as generating content for students to evaluate and reflect on. If any form of cheating or fraud is detected in assignments, research projects or writing pieces, those students must face the consequences of their actions. Some participants proposed harsh punishment, such as expulsion from all academic activities, forfeiting of grades or deregistration from courses or qualifications to avoid cheating or academic dishonesty. Participants proposed, for example, creating case studies, problem-based projects and writing reflections to minimise the use of GenAI apps to write on behalf of the student. Based on measures to combat or prevent academic fraud or dishonesty, MTL mentioned: *"I set problem-based learning tasks for students to reflect, resolve and evaluate each task either or as a group. Each group evaluates the task performed by another group and grades each task according to assessment criteria"*. Moreover, participants are of the view that lecturers have the means to combat cheating by switching from traditional to alternative or authentic assessment. SL1: *"I developed project-based learning as an alternative assessment opportunity for each student to reflect, plan, implement and present the final project as part of the portfolio as an authentic assessment. These alternative assessment tasks minimise the use of ChatGPT to write for the student"*.

### **Empower students with digital literacy skills in using GenAI detection tools as a preventive measure to discourage cheating**

Participants regarded it as their responsibility to set an example by educating students and creating an awareness of the different detector tools available to discourage cheating. One participant empowered her students by exposing them to the use of detector tools to combat cheating in the course. She said: *"Since ChatGPT was launched, I have exposed my students to this generative tool. This was my way to create an awareness that detection tools will be used as a preventive measure to discourage cheating"* (SL2). Another lecturer echoed the same sentiment: *"Internal emails to staff and students posted or sent several communications (social media platforms) that the use of generative software was not allowed. This is a measure to prevent academic dishonesty"* (FP1).

### **Advocate and inculcate principles of integrity, morality, and ethical responsibility in using GenAI-chatbots in academic writing**

Different types of detector tools are available to prevent plagiarism or academic dishonesty (Awan, 2023; Chaka

2023; Lim, 2023). These studies reported that students confirmed using GenAI software to cowrite assignments and other projects. To remedy this challenge, participants should foster academic integrity to improve writing skills and guide them appropriately in the use of AI technologies. Established guidelines in the use of GenAI technologies, for example, citing and referencing in academic writing, should be shared with students. On the other hand, studies revealed an increase in cheating, which raises ethical concerns and moral dilemmas in academia. In most cases, reliance on GenAI can undermine student learning and diminish programme accreditation (Jarrah et al., 2023, Gao et al., 2023).

For some participants, the best way is to advocate strategies to prevent cheating. They believe it is important to advocate principles of integrity and ethical responsibility in using GenAI as a value-driven opportunity to advance their learning. It is vital to expose students to detector software to discourage cheating. AP1 said: *"Before students submit an assessment task, a generative AI declaration is assigned. I explained the similarity in the text of each work. There is no excuse for being ignorant, and if detected cheating, it is punishable with deregistration in my course"*. However, some participants were positive and believed that generative tools should be embraced. SL1 said: *"Let us be positive about generative tools. If the human spirit overcame two world wars, 4IR and the pandemic, it could propel itself beyond the 21st Century. It needs a spirit of 'embrace it, design around it with a possibility attitude' by embracing all the generative AI tools in advancing education."* Some view ethics and collective responsibility as vital strategies to advocate the use of chatbots in teaching and learning. MTL mentioned: *"Inculcate a sense of ethical and collective responsibility amongst our students. I like the critical conversations about ethics and zero-tolerance of fraud using generative AI tools. I am of the view that we should protect the image, qualifications, and values of our institution"*.

### **Design authentic assessment tasks of learning experiences that cannot easily be generated by GenAI chatbots but are applied to real-life contexts for meaningful learning**

Participants opined that authentic learning experiences should be designed and adopted to raise awareness of and prevent cheating among students. Such measures should change the usefulness of chatbots. Webinars on generative AI conversational tools are valuable strategies to increase awareness. Discussion forums and online panels should be created for critical conversations about detector software, and the need for the adoption of detector tools in teaching and learning should be emphasised. One participant echoed:

*"Our college hosted several webinars on ChatGPT and other generative tools. To prevent cheating with generative tools, design and conduct independent research projects that require students to design data collection tools and generate findings or write context-based case studies that need specific solutions that cannot be generated by AI tools"* (SL2).

SL1 said how he created strategies for meaningful and authentic learning opportunities:



"In one of the study units that focuses on topics such as unemployment and poverty as contemporary economic issues, students are grouped, to solve real-world problems, which requires critical thinking, collaboration, application of research skills and creativity for a group presentation is an alternative assessment practice strategy".

## Discussion

This study is underpinned by the Conversation Theory of Pask and methodologically aligned to the constructivist-interpretivist paradigm, exploring academics' views of GenAI as a LLM chatbot. Furthermore, strategies are proposed for the design and application of authentic assessment tasks in real-life contexts as measures to combat academic dishonesty using GenAI software tools in an institution of higher education. Findings revealed that academics are of the view that GenAI tools could be advantageous in teaching and learning, but students must be empowered with digital literacy skills to use GenAI detector tools effectively to combat cheating, fraud, dishonesty, and plagiarism.

The first research question explored participants' view of GenAI, specifically ChatGPT as a conversational LLM tool, to promote and advance the teaching and learning experiences of students. Participants are of the view that the ChatGPT as a GenAI-conversational tool could be used by them (lecturers) to empower students by prompting chatbots to generate specific case studies, problem-solving activities and project-based learning tasks, and advance academic writing skills (Kaplan-Rakowski et al., 2023; Wodecki, 2023; Gamage et al., 2023). Participants echoed sentiments about the usefulness of ChatGPT as one of the generative AI tools in promoting and advancing pedagogy as well as student learning experiences (Megahed et al., 2023; Castillo et al., 2023). FP1 said: *"Based on the usefulness, a participant found the ChatGPT useful and capacitated postgraduate students with academic writing skills"*. Another issue that emerged from the interviews is that disruptive generative AI conversational tools are likely to affect higher education. Mohamed (2023) reported that the lecturers viewed ChatGPT as effective in teaching English foreign language students. However, participants were worried about the affordability and accessibility of these generative AI tools for disadvantaged students (Farrelly & Baker, 2023) because tech companies had invested hugely in gaining an advantage in AI-generative apps for profit-making.

A major issue that participants raised and that requires urgent attention, is ethics. This issue needs to be addressed and incorporated into learning programmes. An awareness of GenAI and ethical considerations in the use of GenAI tools, as well as following a zero-tolerance policy, must be advocated. Participants rated awareness and ethical considerations as very important considerations before adopting chatbots in practice. Furthermore, academics needed to be cognisant of the speed of the emergence of new generative AI tools since the launch of ChatGPT 3.5 in 2022. Studies concurred that ethical compliance in the use of GenAI tools by students is a crucial principle (Baek & Kim, 2023; Eke, 2023; Cotton et al., 2023). These studies reported that lecturers expressed concerns about an increase in academic dishonesty and

cheating by students, which may compromise or jeopardise their academic careers. Moreover, participants agreed that higher education needed to address this concern through policy changes. One participant echoed that *"the chatbot in the cloud [sic...meaning the elephant in the room] is a serious matter and must be dealt with speed to stop cheating and dishonesty by students"* (MTL). Several studies highlighted ethical issues as a great concern which must be addressed urgently (Cooper, 2023; Mhlanga, 2023; Vargas-Murillo et al., 2023). Participants concurred that ethics is at the heart of preventing and detecting cheating or academic dishonesty. SL2 raised: *"Since it became known, generative AI tools sparked heated debates around ethical issues like academic dishonesty and cheating by both academics and students"*. Although AI-detection software is available to detect academic dishonesty and cheating, it is often costly (Awan, 2023; Chaka 2023; Lim, 2023; Oravec, 2023).

The second research question investigated the measures or strategies academics applied to prevent academic cheating by students using ChatGPT in teaching, learning and assessment. Participants proposed authentic assessment or alternative assessment tasks to be applied in real-life contexts, empowering students with digital literacy skills in using GenAI detection tools, and inculcating principles of integrity, morality and ethical responsibility in using GenAI-chatbot strategies as value-added tools to prevent cheating and plagiarism. However, literature as well as participants concur that disruptive generative AI conversational tools are here to stay and will change the face of higher education. The availability of GenAI tools compelled universities to address policy revisions, and awareness amongst staff and students is of vital importance to prevent cheating (Chan, 2023; Dwivedi et al., 2023; Ali, 2023; Ali & Djalilian, 2023; Brendel et al., 2021). According to participants, these are very important strategies against academic dishonesty. To combat academic fraud, cheating and dishonesty, participants felt it was necessary to create and adopt a "zero-tolerance" policy to prevent academic dishonesty. Furthermore, lecturers could design case studies, research- and problem-based projects for students to present the results and have them write and report as a group (Firaina & Sulisworo, 2023; Ismail et al., 2023; Hassoulas et al., 2023). SL2 said, *"In my course, I designed project-based learning opportunities, for engagement, collaboration, implementation, and reporting the final project as part of a portfolio [authentic assessment]"*. Studies advocate the use of AI chatbot strategies to prevent academic dishonesty (Firat, 2023; Ali, 2023; Debby et al., 2023). Several participants believe it is vital to expose students to detector software to discourage cheating. Participants opined that authentic learning experiences should be designed and adopted to raise awareness of and prevent cheating among students. Such assessments should be based on real-life applications to evaluate student competence. MTL echoed: *"Webinars on generative AI conversational tools are a valuable strategy to increase awareness"*. According to Castillo et al. (2023), discussion forums as authentic assessments can create critical conversations about GenAI detector software and motivate students to adopt detector tools.

## Conclusion

Scholarly works have been published since the launch of the first GenAI conversational tool, ChatGPT. The hype around GenAI tools as a large language model (LLM, GenAI) has highlighted a fundamental pedagogical change in higher education. This exploratory study revealed that the ChatGPT phenomenon has affected teaching and learning at a faculty of education. Participants are of the view that GenAI-based tools could advance teaching and learning at the faculty of education. Therefore, it is an opportunity to create teaching and learning for students. Literature and empirical findings in this study reported two critical issues that faculties of education must consider: creating an awareness campaign and developing ethical guidelines for using GenAI tools like ChatGPT by students and academics. Some participants viewed ChatGPT as both a friend and foe for the academic project. To prevent unethical behaviour of students, it is necessary to create and adopt measures to prevent academic dishonesty, cheating or plagiarism. Participants suggest using specific case studies, problem-solving activities, project-based learning tasks and advancing academic writing skills as strategies to empower students.

## Recommendations

This study made significant contributions to advance GenAI-based LLM research in the context of a faculty of education. Furthermore, it advances the growing body of knowledge of generative artificial intelligence applied by academics in education. The study extends the epistemological (knowledge of the subject) of GenAI, in particular, how the Conversation Theory foregrounds the study. Moreover, this study proposed strategies to create an awareness of detector tools to prevent academic dishonesty and protect academic integrity. It is recommended that the universities revise policies, expose students to GenAI detector tools, and create an awareness of ethical considerations for using GenAI tools. Because of the "silence" in the use of GenAI tools for academic purposes, ethics and integrity policies must be adjusted to address this matter urgently. Further research is needed to build on recent gains in academic awareness of GenAI tools for teaching and learning. In addition, research must explore whether ChatGPT as an LLM application enhances students' creativity and critical thinking skills. Finally, it is suggested that ethics, tuition, research, assessment practices, and continuous professional development initiatives in faculties of education be revised.

## Limitations

Scholarly works have been published about the impact, benefits, and challenges of ChatGPT in different fields, subjects, disciplines, and contexts, but to a lesser extent in the faculty of education. Although the hype around GenAI tools as a large language model (LLM) has highlighted the learning opportunities, they also present a threat to academic honesty in higher education. The views of academics from a single institution of higher learning were presented, but further research must be conducted at other HEIs. The findings of this small sampled exploratory study cannot be

generalised; thus, a larger study must be undertaken that may yield different results.

## Acknowledgements

I acknowledge the ethics committee for the approval of ethics to conduct the study and the participants who voluntarily participated in the study. No financial support was received from any funder, and the university's professional editing department edited the manuscript.

## References

- Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions. *Journal of Cleaner Production*, 140527. <https://doi.org/10.1016/j.jclepro.2023.140527>
- Adarkwah, M. A., Amponsah, S., van Wyk, M. M., Huang, R., Tlili, A., Shehata, B., ... & Wang, H. (2023). Awareness and acceptance of ChatGPT as a generative conversational AI for transforming education by Ghanaian academics: A two-phase study. *Journal of Applied Learning and Teaching*, 6(2), 78-93. <https://doi.org/10.37074/jalt.2023.6.2.26>
- Al Matari, A. S., Mukit, A., Al Saadi, S., Balushi, W. A., & Al-Abri, J. S. (2023). Artificial intelligence and the future of teaching in higher education at A'Sharqiyah University (ASU) in Oman. *Proceeding International Pelita Bangsa*, 1(01), 182-200.
- Ali, F. (2023). Let the devil speak for itself: Should ChatGPT be allowed or banned in hospitality and tourism schools? *Journal of Global Hospitality and Tourism*, 2(1), 1-6. <https://digitalcommons.usf.edu/cgi/viewcontent.cgi?article=1026&context=jght>
- Ali, M. J., & Djalilian, A. (2023). Readership awareness series—Paper 4: Chatbots and chatbot-ethical considerations in scientific publications. In *Seminars in ophthalmology* (pp. 1-2). Taylor & Francis.
- Antonenko, P., & Abramowitz, B. (2023). In-service teachers' (mis)conceptions of artificial intelligence in K-12 science education. *Journal of Research on Technology in Education*, 55(1), 64-78. <https://doi.org/10.1080/15391523.2022.2119450>
- Asimov, I. (1955). *The complete robot*. Voyager.
- Awan, A. A. (2023). *Top 10 tools for detecting ChatGPT, GPT-4, Bard, and Claude*. KDnuggets. <https://www.kdnuggets.com/2023/05/top-10-tools-detecting-chatgpt-gpt4-bard-llms.html>
- Baek, T. H., & Kim, M. (2023). Is ChatGPT scary good? How user motivations affect creepiness and trust in generative artificial intelligence. *Telematics and Informatics*, 83, 102030. <https://doi.org/10.1016/j.tele.2023.102030>

- Baidoo-Anu, D., & Owusu Ansah, L. (2023). *Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning*. <https://dx.doi.org/10.2139/ssrn.4337484>
- Bozkurt, A., Xiao, J., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., ... & Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, 18(1), 53-130. <https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/709>
- Brendel, A. B., Mirbabaie, M., Lembcke, T.-B., & Hofeditz, L. (2021). Ethical management of artificial intelligence. *Sustainability*, 13(4), 1974. <https://doi.org/10.3390/su13041974>
- Browne, R. (2023). All you need to know about ChatGPT, the A.I. chatbot that's got the world talking and tech giants clashing. *CNBC*. <https://www.cnbc.com/2023/02/08/what-is-chatgpt-viral-ai-chatbot-at-heart-of-microsoft-google-fight.html>
- Castillo, A. G. R., Silva, G. J. S., Arocutipa, J. P. F., Berrios, H. Q., Rodriguez, M. A. M., Reyes, G. Y., ... & Arias-González, J. L. (2023). Effect of Chat GPT on the digitized learning process of university students. *Journal of Namibian Studies: History Politics Culture*, 33, 1-15. <http://dx.doi.org/10.59670/jns.v33i.411>
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, 6(2), 94-104. <http://dx.doi.org/10.37074/jalt.2023.6.2.12>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 2023(20), 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Creswell, J. W. (2012). *Educational research*. New York.
- Cooper, G. (2023). Examining science education in chatbot: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32(3), 444-452. <https://doi.org/10.1007/s10956-023-10039-y>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12. <https://doi.org/10.1080/14703297.2023.2190148>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eke, D. O. (2023). ChatGPT and the rise of generative AI: A threat to academic integrity? *Journal of Responsible Technology*, 13, 100060. <https://doi.org/10.1016/j.jrt.2023.100060>
- Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Educational Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>
- Firaina, R., & Sulisworo, D. (2023). Exploring the usage of ChatGPT in higher education: Frequency and impact on productivity. *Buletin Edukasi Indonesia*, 2(01), 39-46. <https://doi.org/10.56741/bei.v2i01.310>
- Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, 6(1), 57-63. <https://doi.org/10.37074/jalt.2023.6.1.22>
- Gamage, K. A., Dehideniya, S. C., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: More opportunities than concerns. *Journal of Applied Learning and Teaching*, 6(2), 358-369. <https://doi.org/10.37074/jalt.2023.6.2.32>
- Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *Digital Medicine*, 6, 75. <https://doi.org/10.1038/s41746-023-00819-6>
- Goralski, M. A., & Tan, T. K. (2023). Artificial intelligence: Poverty alleviation, healthcare, education, and reduced inequalities in a post-COVID world. In F. Mazzi, & L. Floridi (Eds.), *The ethics of artificial intelligence for the sustainable development goals* (pp. 97-113). Springer International Publishing. [https://doi.org/10.1007/978-3-031-21147-8\\_6](https://doi.org/10.1007/978-3-031-21147-8_6)
- Griffith, E., & Metz, C. (2023). 'Let 1,000 flowers bloom': A.I. funding frenzy escalates. *The New York Times*. <https://www.nytimes.com/2023/03/14/technology/ai-funding-boom.html>
- Hassoulas, A., Powell, N., Roberts, L., Umla-Runge, K., Gray, L., & Coffey, M. (2023). Investigating marker accuracy in differentiating between university scripts written by students and those produced using ChatGPT. *Journal of Applied Learning & Teaching*, 6(2), 71-77. <https://doi.org/10.37074/jalt.2023.6.2.13>
- Holmes, W., Bialik, M., and Fadel, C. (2019). *Artificial Intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Huallpa, J. J. (2023). Exploring the ethical considerations of using Chat GPT in university education. *Periodicals of Engineering and Natural Sciences*, 11(4), 105-115. <http://>

dx.doi.org/10.21533/pen.v11i4.3770

Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>

Hu, W. (2023). *ChatGPT sparks AI 'gold rush' in Silicon Valley*. Mint (ePaper): California. <https://www.industryweek.com/technology-and-iiot/article/21260367/chatgpt-sparks-ai-gold-rush-in-silicon-valley>

Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, 6(2), 56-63. <https://doi.org/10.37074/jalt.2023.6.2.34>

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say. *Online Journal of Communication and Media Technologies*, 13(4), e202346. <https://doi.org/10.30935/ojcm/13572>

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, 15(16), 12451. <https://doi.org/10.3390/su151612451>

Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research*, 34(2), 313-338. <https://www.learntechlib.org/primary/p/222363/>.

Kelly, A., Sullivan, M., & Strampel, K. (2023). Generative artificial intelligence: University student awareness, experience, and confidence in use across disciplines. *Journal of University Teaching & Learning Practice*, 20(6). <https://doi.org/10.53761/1.20.6.12>

Kumar, R., & Mindzak, M. (2024). Who wrote this? Detecting artificial intelligence-generated text from human-written text. *Canadian Perspectives on Academic Integrity*, 7(1). <http://dx.doi.org/10.55016/ojs/cpai.v7i1.77675>

Li, K. C., & Wong, B. T. M. (2023). Artificial intelligence in personalised learning: A bibliometric analysis. *Interactive Technology and Smart Education*, 20(3). <https://doi.org/10.1108/itse-01-2023-0007>

Lim, H. (2023). *5 content detection tools to tell if content is written by ChatGPT*. HongKiat. <https://www.hongkiat.com/blog/chatgpt-content-detection-tools/>

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. <https://doi.org/10.1016/j.ijme.2023.100790>

Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1), 1-8. <https://doi.org/10.14742/ajet.8695>

McDonald, N., Johri, A., Ali, A., & Hingle, A. (2024). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers and Society: Artificial Intelligence*. <https://doi.org/10.48550/arXiv.2402.01659>

Megahed, F. M., Chen, Y. J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2023). How generative AI models such as ChatGPT can be (mis) used in SPC practice, education, and research? An exploratory study. *Quality Engineering*, 1-29. <http://dx.doi.org/10.48550/arXiv.2302.10916>

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. Education, the responsible and ethical use of ChatGPT towards lifelong learning. *SSRN Electronic Journal*. <http://dx.doi.org/10.2139/ssrn.4354422>

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>

Mijwil, M. M., Hiran, K. K., Doshi, R., Dadhich, M., Al-Mistarehi, A.-H., & Bala, I. (2023). ChatGPT and the future of academic integrity in the artificial intelligence era: A new frontier. *Al-Salam Journal for Engineering and Technology*, 2(2), 116-127. <https://doi.org/10.55145/ajest.2023.02.02.015>

Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, 1-23. <https://doi.org/10.1007/s10639-023-11917-z>

Mondal, S., Das, S., & Vrana, V. G. (2023). How to bell the cat? A theoretical review of generative artificial intelligence towards digital disruption in all walks of life. *Technologies*, 11(2), 44. <https://doi.org/10.3390/technologies11020044>

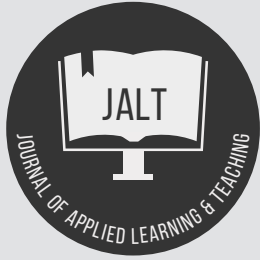
Nemorin, S., Vlachidis, A., Ayerakwa, H. M., & Panagiotis, A. (2023). AI hyped? A horizon scan of discourse on Artificial Intelligence in Education (AIED) and Development. *Learning, Media and Technology*, 48(1), 38-51. <https://doi.org/10.1080/17439884.2022.2095568>

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1). <http://dx.doi.org/10.1177/1609406917733847>

Oravec, J. A. (2023). Artificial intelligence implications for academic cheating: Expanding the dimensions of responsible human-AI collaboration with ChatGPT. *Journal of Interactive Learning Research*, 34(2), 213-237. <https://www.learntechlib.org/primary/p/222340/>

- Oxford Analytica. (2023). *GenAI will transform workplace tasks across industries*. Emerald Expert Briefings. <https://doi.org/10.1108/OXAN-DB279960>
- Pask, G. (1966). *A cybernetic model for some types of learning and mentation*. Bionics Symposium, Dayton, Ohio.
- Pask, G. (1976). *Conversation theory: Applications in education and epistemology*. Elsevier.
- Pask, G., & Scott, B. C. E. (1972). Learning strategies and individual competence. *International Journal of Man-Machine Studies*, 4(3), 217–253. [https://doi.org/10.1016/S0020-7373\(72\)80004-X](https://doi.org/10.1016/S0020-7373(72)80004-X)
- Peres, R., Schreier, M., Schweidel, D., & Sorescu, A. (2023). On ChatGPT and beyond: How generative artificial intelligence may affect research, teaching, and practice. *International Journal of Research in Marketing*. <https://doi.org/10.1016/j.ijresmar.2023.03.001>
- Pisica, A. I., Edu, T., Zaharia, R. M., & Zaharia, R. (2023). Implementing artificial intelligence in higher education: Pros and cons from the perspectives of academics. *Societies*, 13(5), 118. <https://doi.org/10.3390/soc13050118>
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 1-13. <https://doi.org/10.1186/s41039-017-0062-8>
- Rose, K., Massey, V., Marshall, B., & Cardon, P. (2023). IS professors' perspectives on AI-assisted programming. *Issues in Information Systems*, 24(2). [https://doi.org/10.48009/2\\_iis\\_2023\\_115](https://doi.org/10.48009/2_iis_2023_115)
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare*, 11(6), 887. <https://doi.org/10.3390/healthcare11060887>
- Smolansky, A., Cram, A., Radulescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023). Educator and student perspectives on the impact of generative AI on assessments in higher education. In *Proceedings of the tenth ACM conference on Learning@Scale* (pp. 378-382). <https://doi.org/10.1145/3573051.3596191>
- Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, 6(3), 355–366. <https://doi.org/10.1177/20965311231168423>
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 5(236), 433–460. <http://dx.doi.org/10.1093/mind/LIX.236.433>
- Van Wyk, M. M., Adarkwah, M. A., & Amponsah, S. (2023). Why all the hype about ChatGPT? Academics' views of a chat-based conversational learning strategy at an open distance e-learning institution. *Open Praxis*, 15(3), pp. 214–225. <https://doi.org/10.55982/openpraxis.15.3.563>
- Vargas-Murillo, A. R., de la Asuncion, I. N. M., & de Jesús Guevara-Soto, F. (2023). Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. *International Journal of Learning, Teaching and Educational Research*, 22(7). <https://doi.org/10.26803/ijlter.22.7.7>
- Wang, S., & Zhang, Y.-D. (2023). Advances and challenges of deep. *Recent Patents on Engineering*, 17(4), e300522205402. <https://dx.doi.org/10.2174/1872212116666220530125230>
- Wodecki, B. (2023, February 4). *UBS: ChatGPT is the fastest-growing app of all time*. AI Business. <https://aibusiness.com/nlp/ubs-chatgpt-is-the-fastest-growing-app-of-all-time>
- Yang, W. (2022). Artificial intelligence education for young children: Why, what, and how in curriculum design and implementation. *Computers and Education: Artificial Intelligence*, 3, 100061. <https://doi.org/10.1016/j.caeai.2022.100061>

Copyright: © 2024. Micheal M Van Wyk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Higher education assessment practice in the era of generative AI tools

---

Bayode Ogunleye <sup>A</sup>	A	<i>School of Architecture, Technology &amp; Engineering, University of Brighton, Brighton BN2 4GJ, United Kingdom</i>
Kudirat Ibilola Zakariyyah <sup>B</sup>	B	<i>School of Architecture, Technology &amp; Engineering, University of Brighton, Brighton BN2 4GJ, United Kingdom</i>
Oluwaseun Ajao <sup>C</sup>	C	<i>Department of Computing &amp; Mathematics, Manchester Metropolitan University, Manchester, M1 5GD, United Kingdom</i>
Olakunle Olayinka <sup>D</sup>	D	<i>Department of Computer Science, University of Sheffield, Sheffield S1 4DP, United Kingdom</i>
Hemlata Sharma <sup>E</sup>	E	<i>Department of Computing, Sheffield Hallam University, Sheffield S1 2NU, United Kingdom</i>

---

### Keywords

---

Artificial intelligence (AI);  
assessment practice;  
ChatGPT;  
Gemini;  
generative AI (GenAI);  
higher education;  
teaching and learning.

### Abstract

---

The higher education (HE) sector benefits every nation's economy and society at large. However, their contributions are challenged by advanced technologies like generative artificial intelligence (GenAI) tools. In this paper, we provide a comprehensive assessment of GenAI tools towards assessment and pedagogic practice and, subsequently, discuss the potential impacts. This study experimented using three assessment instruments from data science, data analytics, and construction management disciplines. Our findings are two-fold: first, the findings revealed that GenAI tools exhibit subject knowledge, problem-solving, analytical, critical thinking, and presentation skills and thus can limit learning when used unethically. Secondly, the design of the assessment of certain disciplines revealed the limitations of the GenAI tools. Based on our findings, we made recommendations on how AI tools can be utilised for teaching and learning in HE.

---

### Correspondence

[b.ogunleye@brighton.ac.uk](mailto:b.ogunleye@brighton.ac.uk) <sup>A</sup>

---

### Article Info

Received 17 February 2024

Received in revised form 10 March 2024

Accepted 20 March 2024

Available online 1 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.28>

## Introduction

The higher education (HE) sector contributes to every nation's economy and benefits society in various ways. For example, HE contributes to political stability, greater social mobility, improved social capital, crime reduction, greater social cohesion, innovation, trust, and tolerance (Brennan et al., 2013). Over the years, HE has been known to prepare individuals for the future by equipping learners with the required skills for employment. HE provides a pathway to specific careers such as law, pharmacy, and medicine (Harvey, 2000). Unfortunately, this sector is now seen to face uncertainty due to the increasing development of advanced technologies such as GenAI tools. There has been an increasing fear about the use of AI and its effect on education (Daun & Brings, 2023). The rise of GenAI has produced innovative systems such as ChatGPT (Brown et al., 2020) and Gemini (formerly known as Bard), which have taken the world by storm. These tools were a result of innovative solutions developed using large language models (LLMs) such as the generative pre-trained transformers (GPT) series developed by OpenAI, generalised autoregressive pre-training for language understanding (XLNet) developed by Google, Salesforce's conditional transformer language models (CTRL), Google's Pathways language model (PaLM), and Meta's large language model meta-AI (LLaMa). The LLMs have been used for various natural language processing tasks such as questioning and answering (Pochiraju et al., 2023), sentiment analysis (Habbat et al., 2022), topic modelling (Ogunleye et al., 2023), cyberbullying detection (Ogunleye & Dharmaraj, 2023), and fake news detection (Caramancion, 2023). As detailed in Table 1 below, the development of LLMs is mainly dominated by a few large organisations, including Google, Meta, and Microsoft/OpenAI. This is due to the very large amount of data (parameters) used to pre-train these models and the significant computational resources required to build them. However, a growing number of these models are now existing as fine-tuned applications in open-source platforms such as Hugging Face and Stable Diffusion.

Table 1. List of popular large language models (LLMs).

Large Language Model	No Parameters Trained on	Released By	Publication
LLaMA	65 billion	Meta	Touvron et al. (2023)
LaMDA	137 billion	Google	Thoppilan et al. (2022)
PaLM	540 billion	Google	Chowdhery et al. (2023)
UL2	20 billion	Google	Tay et al. (2022)
Codex	12 billion	Microsoft/OpenAI	Chen et al. (2021)
GPT-3	175 billion	Microsoft/OpenAI	Brown et al. (2020)

It is worth mentioning that the parameters of recent LLMs like GPT-4 are undisclosed due to the competitive landscape and the safety implications of large-scale models (OpenAI, 2023a). In academia, Daun and Brings (2023) discussed the use of GenAI tools such as ChatGPT for teaching and learning. The authors inferred that the tools can be used for self-assessment of one's own solution, answering student queries, and generating exercises. Baidoo-Anu and Ansah (2023) added that the systems can support teaching and learning

by providing personalised tutoring, language translation, interactive learning, and automated essay grading. While GenAI solutions can serve as a valuable resource, the technology can also be misused. For example, students can make use of the system for cheating (Cotton et al., 2024); thus several concerns have been raised about these AI tools (Chaudhry et al., 2023; Cotton et al., 2024; Farrokhnia et al., 2023; Halaweh, 2023; Rasul et al., 2023). Specifically, Rasul et al. (2023) and Nikolic et al. (2023) stated that academic integrity is potentially compromised as ChatGPT has proven its competence in achieving success on medical licensing and law exams as well as producing research abstracts/contents, statistical analyses, and computer programs that are not detectable. Some recent attempts at detecting AI-generated content include LLM watermarking models (Sadasivan et al., 2023; Tang et al., 2023) and ChatGPT Checker (OpenAI, 2023b). However, studies like Chaka (2023, 2024) show that AI content detectors are inconsistent and unreliable at identifying AI-generated content. It is increasingly shown that these content detection approaches are failing in their efforts, as the GenAI models that created them are also becoming more sophisticated (*The Times*, 2023). It has been found that fine-tuning these foundational AI models, such as Bidirectional Encoder Representations from Transformers (Devlin et al., 2018) and LLaMA (Touvron et al., 2023) on domain-specific data, makes them even more effective at creating human-like responses with closely aligning domain-relevant contexts such as BioBERT (Lee et al., 2020). This further necessitates the need, especially for HE stakeholders, to measure the impact and define ways to incorporate the AI system for beneficial use in academia.

The impact of GenAI tools on teaching, learning and assessment practice in HE is a hotly debated topic (Rudolph et al., 2023a, 2023b). The use of authentic assessment is widespread in UK universities, and the adoption of it has been praised over traditional forms of assessment as it appears to have reduced the negative effect of class size on student attainment (Richardson, 2015). However, assessment practices face more challenges due to the development of AI tools. Based on this background, few studies evaluated the capabilities of GenAI tools on assessment instruments such as examination questions, essays, and coursework. For example, Mahon et al. (2023) assessed the capabilities of ChatGPT on computer science A-level examination. Finnie-Ansley et al. (2022) evaluated the performance of OpenAI Codex on introductory programming (Python) exams. Furthermore, Bartoli et al. (2024) assessed the performance of ChatGPT on neurosurgical residents' written exams. In summary, few studies have investigated the capacity of ChatGPT by assessing the content (assessment solutions) generated. However, this is limited to medical education and programming context. There is a paucity of studies across STEM disciplines that assessed the performance of GenAI tools on assessment in the HE pedagogic practice. Thus, our study intends to fill this gap. Our work did not aim at making judgments on the use of AI-generated content in academia. However, we aim to assess the performance of GenAI tools in STEM-related disciplines to understand their potential impact on students' learning and development.

Most GenAI systems are still in an early stage of development, and there are limited studies on how solutions generated from these systems can impact assessment practice, such as assessment by coursework or essay. Thus, this paper contributes in several ways to teaching, learning and assessment practice in HE, and these contributions can be summarised as follows.

- This study demonstrates a methodology for evaluating the impact of GenAI tools towards assessment practices in HE. We show that AI tools present an opportunity to evaluate, critique and contextualise information. However, this may appear to be discipline-dependent.
- Our study shows the necessity of modifying curricula to accommodate the evolving skill sets demanded. Additionally, we emphasise the importance of providing proper guidance for the utilisation of GenAI. This entails integrating it into courses under proper supervision, rather than allowing unregulated student use, which could impact their learning experience detrimentally.
- Our findings demonstrate the potential impact of GenAI tools on students' learning when used for assessments unethically. Subsequently, we present assessment instruments and criteria which are helpful for future research and learning support in this era of AI-generated content.

The rest of the paper is organised as follows: Section 2 reviews the literature to provide background knowledge for this study. Section 3 centres on the methods, while Sections 4 and 5 present the results, conclusions, and recommendations.

## Related work

The HE sector has been faced with challenges since the advent of Google's search service in 1997 (Brophy & Bawden, 2005). HE faculty members were overwhelmed with the fear of the tool taking their positions as learners could easily search for materials to learn online through these search engines. However, over the years, this has proven not to be the case. Search engines have been used for teaching, learning, and research effectively, and the educational sector has been one of the sectors that benefitted most from the technology. The recent GenAI tools are conversational agents that can generate human-like content such as texts, images, and videos. For illustration, GenAI applications such as ChatGPT are a form of question-answering models (Wu et al., 2023). In their development, they would usually require example questions, which are labelled datasets in diverse semantic and structural formats, sometimes domain-specific, for the GenAI applications to be pre-trained such that they can give the user the correct answers when prompted (asked a question). There are benchmark datasets that exist and have been used for this purpose; these include arithmetic-type questions (Cobbe et al., 2021) and common sense-type questions (Talmor et al., 2018). Table 2 below provides a summary of some of the benchmark datasets. It is worth

Table 2. List of benchmark datasets for pre-training LLMs.

Reasoning Dataset	Category	Description	Publication
GSM8K	Arithmetic	Math Word Problems (8.5K)	Cobbe et al. (2021)
SVAMP		Math Word Problems (1.0K)	Patel et al. (2021)
ASDiv		Math Word Problems (1.2K)	Miao et al. (2021)
AQuA		Algebraic Word Problems (100K)	Ling et al. (2017)
MAWPS		Math Word Problems (2.4K)	Koncel-Kedziorski et al. (2016)
CSQA	Common Sense	Complex semantics (12.2K)	Talmor et al. (2018)
StrategyQA		Multi-hop strategy (2.8K)	Geva et al. (2021)
BIG-bench (Date)		Inferring Date from context (204)	Srivastava et al. (2022)
BIG-bench (Sports)		Plausibility of sports sentences	
SayCan		Mapping instructions to robot actions (101)	Ahn et al. (2022)

stating that there are other labelled datasets that were used for training the recent LLMs like GPT-4. The datasets are bound to be much larger; however, information about them has not been disclosed. In academia, there are fears about GenAI's effect on HE (Daun & Brings, 2023). Kaplan-Rakowski et al. (2023) investigated teachers' perspectives on the use of GenAI for teaching and learning, using 147 diverse groups of teachers via an online survey. Their survey questions were around technology integration (in terms of awareness, learning, understanding, familiarity, adaptation, and application), participants' perceptions of GenAI implementation in education, and the use of GenAI for teaching. Their result showed that teachers have a positive perception of the use of AI tools. However, that does not translate into actions. Grassini (2023) discussed both the potentials and challenges associated with the integration of generated AI tools in academia. They identified key potentials of AI as able to assist in providing feedback and developing learning materials. Equally, key challenges, such as bias, hallucinations, academic integrity, and data privacy, were highlighted.

Assessment instruments like examinations or coursework are tools which are used to evaluate and enhance student learning. Assessment strongly influences students' learning (Bloxham, 2015). The marking of students' assessments involves evaluating various aspects of students' performance. Some of the elements of assessment include accuracy and validity, the demonstration of learning, the transfer of knowledge, collaboration, and metacognition (Ashford-Rowe et al., 2014). Kim et al. (2019) opined that higher-order skills that cover complex thinking, communication, collaboration, and creativity, which are also referred to as the 4Cs, are the most significant authenticity or future skills criteria. Crawford et al. (2023) emphasised on demonstrating comprehension of a subject to solve complex problems as an assessment criterion rather than regurgitating theories in a textbook. Academics have used criterion-referenced assessment (CRA) successfully based on its reliability, validity, and transparency in assessing learning (Burton, 2006; Liao, 2022; Lok et al., 2016). Assessment criteria play a vital role in defining and assessing students' performance in various educational setups. The criteria offer clear, objective standards against which students' work is evaluated. The criteria must be mentioned clearly; this will help students understand what is required and how the assessment will be marked (Popham, 1997). The approach helps reduce



variations in marks awarded to students. Despite the importance of assessment to learning, the availability of GenAI tools poses several threats to authentic assessment. A major concern is the student usage for cheating. This has brought a lot of attention to HE to implement policy and guidance on the usage of AI systems. Most importantly, to integrate GenAI into academia. Currently, there is no agreed guideline for the usage of GenAI systems in HE, and thus, it is worth assessing the capabilities of the systems across several disciplines to inform policy-making processes. Past studies assessed the performance of GenAI tools using assessment instruments like examination questions and coursework. For example, Malinka et al. (2023) tested ChatGPT's performance on programming tasks and concluded that ChatGPT has the capacity to pass the courses required for a university degree in IT security. Kolade et al. (2024) deployed ChatGPT to generate academic essays on the digital transformation of the health sectors in the global South with suggestions on improving digitally enabled healthcare delivery. Their study showed that ChatGPT 3.5 generates original high-quality content that is hard to distinguish from human-generated content. To conclude, we present a summary of existing key literature that assessed the performance of GenAI tools for pedagogy practice in Table 3. This is beneficial for academics to understand solutions that AI systems can generate in their discipline.

Table 3. Summary of key literature that assessed the capabilities of GenAI tools.

Study	Aim	Method	Findings	Conclusion & Recommendation
Thibaut et al. (2024)	The paper assessed ChatGPT and Bard's abilities to pass the first part of the European Board of Hand Surgery (EBHS) diploma examination.	Prompts and answers from ChatGPT and Bard	The GenAI results were poor. The study showed there is no significant difference in the performance of the GenAI tools when answering the multiple-choice exam questions. However, they are still developing their learning capability.	Their study concluded that ChatGPT and Bard are not capable of passing the first part of the EBHS diploma exam in their current state.
Cuthbert & Simpson (2023)	The study assessed whether ChatGPT could pass Section 1 of the Fellowship of the Royal College of Surgeons (FRCS) examination in Trauma and Orthopaedic Surgery.	Prompts and answers from ChatGPT	ChatGPT scores were 30-35% lower than the FRCS pass rate, with 8.2% lower than the mean score achieved by human candidates of all training levels. ChatGPT had 53.3% in basic science but 0% in trauma.	ChatGPT lacks the higher-order judgment and complex thinking required to pass the FRCS examination. In addition, it also fails to recognise its own limitations.  ChatGPT's deficiencies should be publicised equally as much as its successes to ensure clinicians remain aware of its fallibility.
Mahon et al. (2023)	Assessed ChatGPT-4 abilities on computer science examinations (UK A-Level and Irish Leaving Certificate).	Prompts and answers from ChatGPT	The results indicated that ChatGPT is capable of achieving very high marks on both examinations. Furthermore, their results showed that the performance differences before and after the knowledge cut-off date (September 2021) are minimal. However, ChatGPT struggled with image and symbol questions. Minimal hallucination occurrences were observed.	The occurrence of hallucinations in answers and a few errors in the solutions provided in questions without images call for concern in examination designs.  In the future, ChatGPT can be adopted to gather information and discuss tactics during an examination, and such an interactive process can be recorded.
Gupta et al. (2023)	The study tested the GPT-4 exploitation as an instrument for plastic surgery graduate medical education by evaluating its performance on the Plastic Surgery Inservice Training Examination (PSITE).	Prompts and answers from ChatGPT	GPT-4 answers were 77.3% accurate (187 out of 242 questions were answered correctly)	GPT-4 possesses excellent accuracy and reliability for plastic surgery resident education, over GPT-3.5.  Academics/Practitioners should utilise ChatGPT to enhance their educational curriculum.
Bartoli et al. (2024)	The paper assessed how ChatGPT performs at both generating questions and answering a neurosurgical resident's written exam.	Prompts and answers from ChatGPT	Though the ChatGPT required an iterative process, it answered all its self-generated questions correctly, and there was no difference in response rate for residents between human-generated and AI-generated questions, which could have been attributed to the lack of clarity of the questions.	AI is a promising and powerful tool, but it should be used for specific medical purposes that need to be further determined. To enhance its versatility, the prompts must be carefully, precisely, and reasonably formulated.

Wang et al. (2023)	The authors evaluated ChatGPT's ability using three medical examination datasets. The testing involved ordering ChatGPT to act as a doctor to answer exam questions, provide patient discharge summaries, and provide diagnoses.	Prompts and answers from ChatGPT	GPT3 performed well in the China National Medical Licensing Examination in Chinese (CNMLE), its English version (ENMLE), and the China National Entrance Examination for Postgraduate-Clinical Medicine Comprehensive Ability (NEEPM) exams with scores of 56%, 76% and 62% respectively. However, it had lesser scores compared to GPT-4, that scored 84%, 86%, and 82%. Nevertheless, in other areas/criteria such as verbal fluency, open and close domain hallucinations, etc., both GPT 3 and GPT-4 took the turns, that is where one takes the lead, and the other takes the tail. In summary, GPT-4 appears more promising.	In general, ChatGPT accuracy was good. However, it is worth stating that the system still struggles with key medical information, insufficient diagnosis and false information.
Rudolph et al. (2023b)	The study compared the performances of ChatGPT, Bing chat and Bard using 15 questions.	Prompts and answers from ChatGPT, Bing chat and Bard	The study showed that ChatGPT-4 performed better than Bing chat and Bard. However, in general, the chatbots are not yet at the A or B level grade. In addition, the study revealed a multi-disciplinary test that is relevant for HE assessments.	Recommends the integration of AI systems to support learning.  Academics should encourage the use of oral exams to assess student learning.  Academics should endeavour to use authentic assessments for assessing student learning outcomes.  Students should declare the use of AI systems in assessments
Flores-Cohaila et al. (2023)	The paper assessed ChatGPT 3.5 and GPT-4 accuracies on the Peruvian National Medical Examination (Examen Nacional de Medicina [ENAM]) and the identification of factors associated with incorrect answers provided by ChatGPT. The study made use of 180 multiple-choice questions from the ENAM 2022 data set.	Prompts and answers from ChatGPT	On the ENAM, GPT-4 achieved an accuracy of 86%, GPT-3.5 77%, and the 1025 examinees 55%, implying that GPT 3 and GPT 4 outperformed human candidates.  Their results suggest incorrect answers were associated with the difficulty of questions, which may resemble human performance.	The study concluded that ChatGPT (GPT-3.5 and GPT-4) can achieve expert-level performance on the ENAM, outperforming most examinees.
Kunitsu (2023)	The study evaluated GPT-4's ability to answer questions from the Japanese National Examination for Pharmacists (JNEP).	Prompts and answers from ChatGPT	GPT-4 had an accuracy score of over 60%.	The authors concluded that though the bot accuracy rate is good, and they (ChatGPT-4) can support pharmacists' capabilities, the bots have limitations, especially in handling highly specialised questions, calculation questions, and questions requiring diagrams.
Parker et al. (2024)	The authors investigated the use of AI in undergraduate assessment with a focus on the ability of graduate teaching assistants (GTAs) to spot AI-generated assessments. They also examined the performance of ChatGPT in producing high-quality work.	ChatGPT	ChatGPT excelled the average student in all classes, with top marks in 8 of the 10 classes while the GTAs were only able to identify 50% of the AI-generated assessments solutions.	AI consistently performed well, indicating its robust capability in handling a wide range of assessment types and academic subjects. They equally concurred that there are variations in the literature on models of results when AI is used in assessment.
Funk et al. (2024)	The paper evaluated GenAI tools in terms of reliability using 450 medical examination questions.	ChatGPT 3.5 and 4	ChatGPT-4 showed better accuracy than ChatGPT-3.5.	Though ChatGPT-4 performance outweighed ChatGPT-3.5, the significance of human factor in medical education and clinical decision making, cannot be overemphasised.

## Methodology

This section provides details on the GenAI, and assessment tools employed. In this paper, we prepared three case studies of assessments from the data science, data analytics, and construction management disciplines (as shown in Appendix A – C). The selection of these disciplines was based on the authors' expertise. OpenAI's GPT-4 and Google's Bard (now Gemini) are well-performing GenAI tools among the LLMs that have transformed how we interact with machines and process massive amounts of text information (Dhanvijay et al., 2023). These models can generate text like human beings and perform a variety of tasks (Mohamadi et al., 2023),

such as text completion, language translation, and content generation. As a result of their versatility, effectiveness, popularity, innovativeness, and usage in the literature, both of these tools are compelling choices for users across a range of industries and research fields. For these reasons, our study employed ChatGPT-4 and Bard as our GenAI tools to attempt the assessments and generate solutions to the tasks. These tools were accessed and used for our experiment within the entire month of September 2023. The assessment case studies in subject areas of data analytics, data science, and construction management were chosen at the master's degree level because all authors have taught at that level of study and are well-informed and equipped to assess work submitted for assessment in those disciplines. We used the standard grading scale (A – G) to provide marks for the solutions generated by the AI tools. The grade scale is selected because it is commonly used, especially in UK universities (University of Aberdeen, n.d.; Zarb et al., 2023). The grade scale can be interpreted as shown in Table 4.

Table 4. Interpretation of grades.

Grade	Interpretation
A	Excellent
B	Very good
C	Good
D	Pass
E	Marginal Fail
F	Fail
G	No submission

Furthermore, each assessment case study made use of criterion-referenced assessment due to the sound theoretical rationale, effectiveness, suitability, appropriateness, and applicability (Liao, 2022; Lok et al., 2016). Our assessment criteria are developed based on criteria/skill sets identified in the literature. Our literature findings suggest comprehension (Crawford et al., 2023; Rudolph et al., 2023b), analysis, and accuracy (Ashford-Rowe et al., 2014; Rudolph et al., 2023b). We added discussion and presentation criteria. The latter criteria were selected to assess critical thinking, coherence, hallucination, and bias. Incidentally, the criteria are common within the three disciplines. Table 5 presents the assessment criteria used for grading the solutions generated by ChatGPT and Bard (also, the marking rubric can be seen in Appendix D). For evaluation purposes, the assessments and marks awarded were moderated by authors and subject-area colleagues. We involved at least two academic tutors for each case study to moderate the assessment instruments and the marks awarded. This is to ensure instructions are clear and the level of difficulty suits the level of a master's degree. In addition, the approach aims to minimise bias and variance in marks awarded.

Table 5. Assessment criteria.

Assessment Criteria	Assessment Skills Outcome	Description	Grade (Scale: A-G)
Comprehension and Application	Subject Knowledge, Problem Solving	Understanding of the problem, explanation, Interpretability, and use of literature from a wide range of disciplines.	
Analysis and Synthesis, Accuracy and Relevance	Innovative/Creativity, Problem Solving	Reasoned debate and analysis, and use of theory relevant to the subject.	
Evaluation/Discussion of Result	Critical Thinking	Critical evaluation demonstrated. Implications of the critique identified.	
Presentation/Readability	Coherence, communication, organisation, independent learning	Coherent use of language, structure, presentation and referencing.	

## Result and discussion

The marks awarded to the solution provided by ChatGPT and Bard for each of the assessments are presented in Table 6. Marks were awarded by authors with subject knowledge and moderated by at least two other academic tutors in the field. The first assessment is a data analytics problem that involves a binary classification task in the context of heart failure prediction to make informed decisions for a patient's survival. The tasks involve data exploration, fitting machine learning (ML) algorithms with imbalanced and balanced dataset, evaluation of the ML algorithms performance using appropriate metrics and lastly, recommendation of the best performing model.

Table 6. Grades awarded.

Assessment Criteria	Assessment Number GenAI Model	1		2		3	
		ChatGPT	Bard	ChatGPT	Bard	ChatGPT	Bard
Comprehension and Application		B	C	C	B	B	D
Analysis and Synthesis, Accuracy and Relevance		A	B	B	B	C	D
Evaluation/Discussion of Result		B	C	C	A	D	E
Presentation/Readability		B	B	C	B	D	D

The solution generated by ChatGPT shows that the system possesses fine-grained interactive features, subject knowledge, problem-solving, analytical, critical thinking, and presentation skills. The report shows a coherent discussion of the given tasks. For example, the system provided the metadata, performed descriptive statistics, and performed data pre-processing, like the missing values was checked, and feature scaling was done. Surprisingly, the system was able to identify the target variable without stating it and identified the number of classes of the target variable (binary in this case). The system applied the oversampling technique to rebalance the minority class. Afterwards, all the ML algorithms were fitted appropriately. The system offered the flexibility to choose the proportion of data for training and testing purposes. Thereafter, the performance of the ML algorithms was evaluated using metrics such as accuracy, precision, recall, and F1-score. In addition, the system produced bar plots of the ML performance result and discussed the importance of the evaluation metrics in detail. Furthermore, the system provided a critical discussion of why a simple and interpretable model is preferred in the medical domain. However, it is worth stating that the system failed to use the "imblearn" library (Python module) to implement

SMOTE (synthetic minority oversampling technique) technique for rebalancing the class distribution. Also, the solution did not provide any supporting justification.

Similarly, the solution generated by Bard indicates that the system 'possesses' subject knowledge, problem-solving, and analytical skills. The report shows some discussion of the given task. For illustration, the system provided the metadata, performed descriptive statistics, and performed data pre-processing, such as the missing values and feature scaling. Also, the system was able to identify the target variable and the number of classes of the target variable. In comparison to the ChatGPT solution, Bard was able to utilise the SMOTE (synthetic minority oversampling technique) technique to rebalance the class distribution of the target variable. The system used 75% of the data for training and 25% for testing purposes. Afterwards, all ML algorithms were fitted, and the performance was evaluated using appropriate metrics such as accuracy, precision, recall, and F1-score. Surprisingly, all models produced from Bard showed better performance than results produced from ChatGPT. In summary, for Assessment 1, both GenAI tools were able to solve the problem provided. However, unlike ChatGPT, Bard provided references, links to source code and materials useful for the analysis/further reading. These findings are consistent with the study of Kolade et al. (2024), which found that ChatGPT struggles with referencing in assessment tasks. Furthermore, Bard appears to be technically better than ChatGPT. However, it is limited in its interactive ability, flexibility, critical thinking, discussions, and presentation of results.

The second assessment task involved a data science problem that required a reasonable amount of critical thinking. The task requirement was to provide multi-class (of three levels) text classification of a high-dimensional cancer dataset. Due to the size of this dataset, only the URL link was given in the task specification. This was presented in the case of a conventional ML task, where students may be required to pull data from some repository or external database to provide the solution to their work. A further requirement was the suggestion of using a transformer model such as BERT to convert the text into an embedding space before classification. Finally, there was the need to offer an open-ended recommendation for improving model performance in the context of providing a biomedical solution.

Although in terms of analysis, synthesis, accuracy, and relevance, Bard tended to have a similar score. ChatGPT gave a theoretical solution and pseudocode for solving the task and a good attempt to explain the steps involved. Bard gave a well-written comprehensive code but relatively smaller snippets of text in between. For evaluation and discussion of results, Bard gave well-detailed evaluation results of precision, recall, f-measure, and accuracy with impressive results attained. However, ChatGPT could not go further to give a realistic solution due to its limited ability to handle big data. Bard, on the other hand, discussed various ways of improving model performance and optimisation, including the use of biomedical data to further fine-tune the transformer model. ChatGPT presented the results of the tasks in a more organised fashion, having well-formatted and numbered sections. However, as the solution for the task

was not completed, significant marks were lost here. Overall, in providing the solution to Assessment 2, it is observed that Bard appeared to have given a comprehensive solution in its implementation, while ChatGPT has provided more context, albeit in a much less robust technical analysis of the problem. In terms of comprehension and application, Bard tended to show greater coverage of the task requirement and more attention to detail. We are unsure if it was due to the nature of the case study being more data-intensive or the use of LLMs, which required significant computational power.

The third assessment is a construction management problem. The solution by ChatGPT is commendable. The information provided showed that the system possesses the subject knowledge, understanding and problem-solving skills and showed innovativeness and creativity in solving the problem. However, one schematic that takes care of all the assessment design criteria, as demanded by the assessment design, would have sufficed. We noted that some important information needed was not provided. For example, the purpose of the meeting for which the slide became a requirement was entirely omitted or ignored. Secondly, the context, that is, the Qukzome Community Health Center, was not considered. In addition, the title, as proposed by the GenAI, failed to take cognisance of the variables or themes mentioned in the assignment design. Consequently, dispute, a key variable in the assessment, was not reckoned in the slide suggested by ChatGPT. This implies that critical analysis, thinking, and evaluation were not taken care of. Similarly, Bard's solution did not capture information on how the one-slide PowerPoint presentation would be delivered, though its detailed explanations can be valuable. However, the absence of information or guidance on the slide presentation, with regards to the Iron Triangle and other constraints, reveals that Bard lacks complete comprehension of the task required, and so limits problem-solving efficiency. Bard failed to show either innovativeness or creativity in solving this problem. There was a complete absence of any schematic that recognises all the assessment design criteria.

The key criteria for critical thinking and evaluation, as omitted by ChatGPT, were equally omitted by Bard. Though Bard gave some interrelationships among the criteria mentioned, the information on the slide and the criteria were poorly presented. This is a clear indication that critical analysis, thinking, and evaluation of the assessment design was completely missing. Thus, it falls below that of ChatGPT, which gave an illustration of the iron triangle. In general, both solutions provided are coherent, organised, and well-communicated. However, the solutions did not emphasise the need to "avoid or reduce dispute", another key factor that buttressed the two AI tools' limitations for problem/project-specific or case-specific assessment. The assessment design did not request that references be provided, so both AI tools were graded equally on this aspect.

Our findings are two-fold. First, overall, the results from two of the assessments evidenced that GenAI tools possess interactive features (communication), subject knowledge, problem-solving, analytical, discipline-dependent critical thinking, research, and presentation skills. These are the essential skills that the assessments aim to measure (Lok

et al., 2016). For the construction management-related assessment, the tools struggled with complex thinking skills. This is an indication that the system performance varies across the disciplines. This is due to the availability of learning resources. The availability of learning resources online varies across disciplines, which impacts the amount of content on which the LLMs were trained. The data science and data analytics disciplines encourage open learning approaches such as open sources, boot camps, and open access. Thus, there are many publicly available learning resources and opportunities. The LLMs were trained with more resources in these disciplines and, thus, can generate more content when asked. This is evidenced by the links provided by Bard. Several links to relevant online solutions, including source codes (for example, GitHub links), were produced by Bard. In practice, this implies that relevant and largely accurate content can be generated in the field of data science and data analytics for teaching and learning with a low level of human instructors for intervention. However, fields like construction management are not yet at that stage.

Considering the current state of teaching and learning in HE, we establish that GenAI tools can limit the development of learning and employability skills when used unethically, specifically in areas of data analytics and data science. This is because solutions provided by the systems showed a fine level of accuracy and relevance. However, the performances of the GenAI tools vary across disciplines, as evidenced by the result from the third assessment. For this class of disciplines, the more project-based the assessment design or practice is, the more difficult it becomes for the GenAI tools. Thus, there is a need for future research to consider experimenting in other disciplines to improve generalisation as our results are limited to the fields of data analytics, data science and construction management. Based on our results, we recommend that tutors urgently consider redesigning the assessment instrument considering the information students can develop from the GenAI tools. Similar to the results of Srivastava et al. (2022), our analysis suggests that when GenAI is used unethically, learning and development of critical thinking skills are hindered.

## Conclusions and recommendations

In this study, we aimed to assess the performance of GenAI tools in STEM-related disciplines to understand their potential impact on students' learning and development. We prepared three case study assessments in the data analytics, data science, and construction management disciplines using ChatGPT and Bard as our GenAI tools. Our results showed that GenAI tools possess subject knowledge, problem-solving, analytical, critical thinking, and presentation skills and, thus, can limit the development of students' learning when used unethically. However, this is discipline-dependent, as two sets of results emerged. It is worth stating that there were minimal occurrences of hallucination and bias in the solutions provided.

In practice, we recommend that the HE sector takes an urgent step in incorporating both GenAI systems for teaching and learning and academic AI content detector features into the plagiarism-detecting system. This is because AI content

detectors (as a stand-alone system) appear to still be in the process of refinement, lacking the capability to differentiate between AI-generated content and human-written text consistently and convincingly (Chaka, 2023, 2024). Thus, our recommendation agrees with Chaka (2024), who proposed the use of both AI content detectors and plagiarism detection tools together with human reviewers. Furthermore, there is a need to re-design assessments. Academic tutors need to get familiar with GenAI systems and thus ensure authentic assessments are prepared to limit students' use of generating solutions from the AI systems. This can be achieved by strategically contextualising assessment. In addition, we encourage the use of presentation as a tool to evidence student learning outcomes. Interestingly, this can be achieved via different formats, such as in-person, virtual, or video recording. Alternatively, we recommend that tutors include AI-generated solutions (including the variants) in the assessment brief, and the assessment task can be in the form of a reflective learning approach. In this case, students can produce a report that critically reflects on the GenAI outputs (solutions) to the assessment task and proffer solutions for optimality.

Similarly, for student engagement, we recommend the use of the GenAI systems as an interactive tool during teaching and learning sessions to stimulate the student learning environment. This can be achieved in various ways. For example, students can engage in a comparative activity (group work exercise) during lectures where they are required to discuss and compare their findings of the AI-generated solutions to a case study or our case study assessments in Appendix A – C (if related). Furthermore, students can critique AI-generated solutions to research essay questions during lecture sessions. Moreover, the use of question banks can be helpful. In this scenario, students can generate questions on a particular topic to assess what they have learnt. Thus, these questions can be used for quick formative quizzes during the lecture session. To conclude, we recommend the use of different assessment tools ranging from in-person class test, which can be conducted online or through a written test, contextualised authentic assessments that are reflection-based, discussion-based assessment, and laboratory-based assessment (where applicable) to improve on assessing student learning outcomes in the GenAI era.

## References

- Ahn, M., Brohan, A., Brown, N., Chebotar, Y., Cortes, O., David, B., Finn, C., Fu, C., Gopalakrishnan, K., Hausman, K., Herzog, A., Ho, D., Hsu, J., Ibarz, J., Ichter, B., Irpan, A., Jang, E., Ruano, R. J., Jeffrey, K., ... Amodei, D. (2022). *Do as I can, not as I say: Grounding language in robotic affordances*. ArXiv Preprint ArXiv:2204.01691.
- Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, 39(2), 205–222. <https://doi.org/10.1080/02602938.2013.819566>
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding

- the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62. <http://dx.doi.org/10.61969/jai.1337500>
- Bartoli, A., May, A. T., Al-Awadhi, A., & Schaller, K. (2024). Probing artificial intelligence in neurosurgical training: ChatGPT takes a neurosurgical residents written exam. *Brain and Spine*, 4, 102715. <https://doi.org/10.1016/j.bas.2023.102715>
- Bloxham, S. (2015). Assessing assessment: New developments in assessment design, feedback practices and marking in higher education. In H. Fry, S. Ketteridge & S. Marshall (Eds.), *A handbook for teaching and learning in higher education: Enhancing academic practice* (4th ed., pp. 107-122). Taylor & Francis (Routledge), Abingdon and New York. <http://insight.cumbria.ac.uk/1579>
- Brennan, J., Durazzi, N., & Sene, T. (2013). *Things we know and don't know about the wider benefits of higher education: A review of the recent literature*. <http://dx.doi.org/10.13140/RG.2.1.1200.0806>
- Brophy, J., & Bawden, D. (2005). Is Google enough? Comparison of an internet search engine with academic library resources. *Aslib Proceedings*, 57(6), 498–512. <https://doi.org/10.1108/00012530510634235>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. <https://doi.org/10.48550/arXiv.2005.14165>
- Burton, K. (2006). Designing criterion-referenced assessment. *Journal of Learning Design*, 1(2), 73–82. <http://dx.doi.org/10.5204/jld.v1i2.19>
- Caramancion, K. M. (2023). Harnessing the power of ChatGPT to decimate mis/disinformation: Using ChatGPT for fake news detection. *2023 IEEE World AI IoT Congress (AlloT)*, 42–46. <https://doi.org/10.1109/AlloT58121.2023.10174450>
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, 6(2), 94-104. <https://doi.org/10.37074/jalt.2023.6.2.12>
- Chaka, C. (2024). Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. *Journal of Applied Learning and Teaching*, 7(1), 1-12. <https://doi.org/10.37074/jalt.2024.7.1.14>
- Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—A case study. *Cogent Education*, 10(1), 2210461. <https://doi.org/10.1080/2331186X.2023.2210461>
- Chen, M., Tworek, J., Jun, H., Yuan, Q., Pinto, H. P. de O., Kaplan, J., Edwards, H., Burda, Y., Joseph, N., Brockman, G., Ray, A., Puri, R., Krueger, G., Petrov, M., Khlaaf, H., Sastry, G., Mishkin, P., Chan, B., Gray, S., ... Zaremba, W. (2021). *Evaluating large language models trained on code*. ArXiv Preprint ArXiv:2107.03374.
- Chowdhery, A., Narang, S., Devlin, J., Bosma, M., Mishra, G., Roberts, A., Barham, P., Chung, H. W., Sutton, C., Gehrmann, S., Schuh, P., Shi, K., Tsyashchenko, S., Maynez, J., Rao, A., Barnes, P., Tay, Y., Shazeer, N., Prabhakaran, V., Reif, E., ... Fiedel, N. (2023). Palm: Scaling language modeling with pathways. *Journal of Machine Learning Research*, 24(240), 1–113. ArXiv Preprint arXiv:2204.02311.
- Cobbe, K., Kosaraju, V., Bavarian, M., Chen, M., Jun, H., Kaiser, L., Plappert, M., Tworek, J., Hilton, J., Nakano, R., Hesse, C., & Schulman, J. (2021). *Training verifiers to solve math word problems*. ArXiv Preprint ArXiv:2110.14168.
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228–239. <https://doi.org/10.1080/14703297.2023.2190148>
- Crawford, J., Cowling, M., & Allen, K.-A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3), 2. <http://dx.doi.org/10.53761/1.20.3.02>
- Cuthbert, R., & Simpson, A. I. (2023). Artificial intelligence in orthopaedics: Can Chat Generative Pre-trained Transformer (ChatGPT) pass Section 1 of the Fellowship of the Royal College of Surgeons (Trauma & Orthopaedics) examination? *Postgraduate Medical Journal*, 99(1176), 1110–1114. <https://doi.org/10.1093/postmj/qgad053>
- Daun, M., & Brings, J. (2023). How ChatGPT will change software engineering education. *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education* (Vol. 1., 110–116). <https://doi.org/10.1145/3587102.3588815>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). *Bert: Pre-training of deep bidirectional transformers for language understanding*. ArXiv Preprint ArXiv:1810.04805.
- Dhanvijay, A. K. D., Pinjar, M. J., Dhokane, N., Sorte, S. R., Kumari, A., Mondal, H., & Dhanvijay, A. K. (2023). Performance of large language models (ChatGPT, Bing Search, and Google Bard) in solving case vignettes in physiology. *Cureus*, 15(8). <https://doi.org/10.7759/cureus.42972>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1–15. <https://doi.org/10.1080/14703297.2023.2195846>
- Finnie-Ansley, J., Denny, P., Becker, B. A., Luxton-Reilly, A., & Prather, J. (2022). The robots are coming: Exploring

- the implications of OpenAI codex on introductory programming. *Proceedings of the 24th Australasian computing education conference* (pp. 10–19). <https://doi.org/10.1145/3511861.3511863>
- Flores-Cohaila, J. A., Garcia-Vicente, A., Vizcarra-Jiménez, S. F., la Cruz-Galán, J. P., Gutiérrez-Arratia, J. D., Torres, B. G. Q., & Taype-Rondan, A. (2023). Performance of ChatGPT on the Peruvian national licensing medical examination: Cross-sectional study. *JMIR Medical Education*, *9*(1), e48039. <http://dx.doi.org/10.2196/48039>
- Funk, P. F., Hoch, C. C., Knoedler, S., Knoedler, L., Cotofana, S., Sofo, G., Bashiri Dezfouli, A., Wollenberg, B., Guntinas-Lichius, O., & Alfertshofer, M. (2024). ChatGPT's response consistency: A study on repeated queries of medical examination questions. *European Journal of Investigation in Health, Psychology and Education*, *14*(3), 657–668. <https://doi.org/10.3390/ejihpe14030043>
- Geva, M., Khashabi, D., Segal, E., Khot, T., Roth, D., & Berant, J. (2021). Did Aristotle use a laptop? A question answering benchmark with implicit reasoning strategies. *Transactions of the Association for Computational Linguistics*, *9*, 346–361. [https://doi.org/10.1162/tacl\\_a\\_00370](https://doi.org/10.1162/tacl_a_00370)
- Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, *13*(7), 692. <https://doi.org/10.3390/educsci13070692>
- Gupta, R., Park, J. B., Herzog, I., Yosufi, N., Mangan, A., Firouzbakht, P. K., & Mailey, B. A. (2023). Applying GPT-4 to the plastic surgery inservice training examination. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, *87*, 78–82. <https://doi.org/10.1016/j.bjps.2023.09.027>
- Habbat, N., Anoun, H., & Hassouni, L. (2022). Combination of GRU and CNN deep learning models for sentiment analysis on French customer reviews using XLNet model. *IEEE Engineering Management Review*, *51*(1), 41–51. <https://doi.org/10.1109/EMR.2022.3208818>
- Halaweh, M. (2023). ChatGPT in education: Strategies for responsible implementation. *Contemporary Educational Technology*, *15*(2). <http://dx.doi.org/10.30935/cedtech/13036>
- Harvey, L. (2000). New realities: The relationship between higher education and employment. *Tertiary Education & Management*, *6*(1), 3–17. <https://doi.org/10.1023/A:1009685205201>
- Kaplan-Rakowski, R., Grotewold, K., Hartwick, P., & Papin, K. (2023). Generative AI and teachers' perspectives on its implementation in education. *Journal of Interactive Learning Research*, *34*(2), 313–338. [https://www.researchgate.net/publication/373821343\\_Generative\\_AI\\_and\\_Teachers'\\_Perspectives\\_on\\_Its\\_Implementation\\_in\\_Education](https://www.researchgate.net/publication/373821343_Generative_AI_and_Teachers'_Perspectives_on_Its_Implementation_in_Education)
- Kim, S., Raza, M., & Seidman, E. (2019). Improving 21st-century teaching skills: The key to effective 21st-century learners. *Research in Comparative and International Education*, *14*(1), 99–117. <https://doi.org/10.1177/1745499919829214>
- Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon*, *10*(4). <https://doi.org/10.1016/j.heliyon.2024.e25953>
- Koncel-Kedziorski, R., Roy, S., Amini, A., Kushman, N., & Hajishirzi, H. (2016). MAWPS: A math word problem repository. *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies* (pp. 1152–1157). <https://doi.org/10.18653/v1/N16-1136>
- Kunitsu, Y. (2023). The potential of GPT-4 as a support tool for pharmacists: Analytical study using the Japanese national examination for pharmacists. *JMIR Medical Education*, *9*, e48452. <https://doi.org/10.2196/48452>
- Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, *36*(4), 1234–1240. ArXiv Preprint arXiv:1901.08746.
- Liao, L. (2022). Dancing with explicit criteria or marginalising them: The complexity of grading student work and the reconstruction of the meaning of criterion-referenced assessment. *Teaching in Higher Education*, 1–16. <https://doi.org/10.1080/13562517.2022.2119076>
- Ling, W., Yogatama, D., Dyer, C., & Blunsom, P. (2017). Program induction by rationale generation: Learning to solve and explain algebraic word problems. ArXiv Preprint ArXiv:1705.04146.
- Lok, B., McNaught, C., & Young, K. (2016). Criterion-referenced and norm-referenced assessments: Compatibility and complementarity. *Assessment & Evaluation in Higher Education*, *41*(3), 450–465. <http://dx.doi.org/10.1080/02602938.2015.1022136>
- Mahon, J., Mac Namee, B., & Becker, B. A. (2023). No more pencils no more books: Capabilities of generative AI on Irish and UK computer science school leaving examinations. *Proceedings of the 2023 conference on United Kingdom & Ireland computing education research* (pp. 1–7). <https://doi.org/10.1145/3610969.3610982>
- Malinka, K., Perešini, M., Firc, A., Hujnák, O., & Janus, F. (2023). On the educational impact of ChatGPT: Is artificial intelligence ready to obtain a university degree? *Proceedings of the 2023 conference on innovation and technology in computer science education* (pp. 47–53). <https://doi.org/10.48550/arXiv.2303.11146>
- Miao, S.-Y., Liang, C.-C., & Su, K.-Y. (2021). A diverse corpus for evaluating and developing English math word problem solvers. ArXiv Preprint ArXiv:2106.15772.
- Mohamadi, S., Mujtaba, G., Le, N., Doretto, G., & Adjero, D. A. (2023). *ChatGPT in the age of generative AI and*

- large language models: A concise survey. ArXiv Preprint ArXiv:2307.04251.
- Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 48(4), 559–614. <https://doi.org/10.1080/03043797.2023.2213169>
- Ogunleye, B., & Dharmaraj, B. (2023). The use of a large language model for cyberbullying detection. *Analytix*, 2(3), 694–707. <https://doi.org/10.3390/analytix2030038>
- Ogunleye, B., Maswera, T., Hirsch, L., Gaudoin, J., & Brunson, T. (2023). Comparison of topic modelling approaches in the banking context. *Applied Sciences*, 13(2), 797. <https://doi.org/10.3390/app13020797>
- OpenAI. (2023a). *GPT-4 Technical report*. <https://doi.org/10.48550/arXiv.2303.08774>
- OpenAI. (2023b). *New AI classifier for indicating AI-written text*. OpenAI. <https://openai.com/blog/new-ai-classifier-for-indicating-ai-written-text>
- Parker, L., Carter, C., Karakas, A., Loper, A. J., & Sokkar, A. (2024). Graduate instructors navigating the AI Frontier: The role of ChatGPT in higher education. *Computers and Education Open*, 6, 100166. <https://doi.org/10.1016/j.caeo.2024.100166>
- Patel, A., Bhattamishra, S., & Goyal, N. (2021). *Are NLP models really able to solve simple math word problems?* ArXiv Preprint ArXiv:2103.07191.
- Pochiraju, D., Chakilam, A., Betham, P., Chimulla, P., & Rao, S. G. (2023). Extractive summarization and multiple choice question generation using XLNet. *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 1001–1005. <https://ieeexplore.ieee.org/document/10142220>
- Popham, W. J. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, 55, 72–75. [https://www.skidmore.edu/assessment/handbook/Popham\\_1997\\_Whats-Wrong\\_and-Whats-Right\\_With-Rubrics.pdf](https://www.skidmore.edu/assessment/handbook/Popham_1997_Whats-Wrong_and-Whats-Right_With-Rubrics.pdf)
- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1), 41-56. <http://dx.doi.org/10.37074/jalt.2023.6.1.29>
- Richardson, J. T. E. (2015). Coursework versus examinations in end-of-module assessment: A literature review. *Assessment & Evaluation in Higher Education*, 40(3), 439–455. <https://doi.org/10.1080/02602938.2014.919628>
- Rudolph, J., Tan, S., & Tan, S. (2023a). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Rudolph, J., Tan, S., & Tan, S. (2023b). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). *Can AI-generated text be reliably detected?* ArXiv Preprint ArXiv:2303.11156.
- Srivastava, A., Rastogi, A., Rao, A., Shoeb, A. A. M., Abid, A., Fisch, A., Brown, A. R., Santoro, A., Gupta, A., Garriga-Alonso, A., Kluska, A., Lewkowycz, A., Agarwal, A., Power, A., Ray, A., Warstadt, A., Kocurek, A. W., Safaya, A., Tarzarv, A., Xiang, A., ... Ippolito, D. (2022). *Beyond the imitation game: Quantifying and extrapolating the capabilities of language models*. ArXiv Preprint ArXiv:2206.04615. <https://doi.org/10.48550/arXiv.2206.04615>
- Talmor, A., Herzig, J., Lourie, N., & Berant, J. (2018). *CommonsenseQA: A question answering challenge targeting commonsense knowledge*. ArXiv Preprint ArXiv:1811.00937. <https://doi.org/10.18653/v1/N19-1421>
- Tang, R., Chuang, Y.-N., & Hu, X. (2023). *The science of detecting LLM-generated texts*. ArXiv Preprint ArXiv:2303.07205. <https://doi.org/10.48550/arXiv.2303.07205>
- Tay, Y., Dehghani, M., Tran, V. Q., Garcia, X., Wei, J., Wang, X., Chung, H. W., Shakeri, S., Bahri, D., Schuster, T., Zheng, H. S., Zhou, D., Houshy, N., & Metzler, D. (2022). *UL2: Unifying language learning paradigms*. ArXiv Preprint ArXiv:2205.05131. <https://doi.org/10.48550/arXiv.2205.05131>
- The Times*. (2023). OpenAI gives up on detection tool for AI-generated text. <https://www.thetimes.co.uk/article/openai-ai-detection-tool-writing-chatgpt-creator-d907jwqkh>
- Thibaut, G., Dabbagh, A., & Liverneaux, P. (2024). Does Google's Bard Chatbot perform better than ChatGPT on the European hand surgery exam?. *International Orthopaedics*, 48(1), 151-158. <https://doi.org/10.1007/s00264-023-06034-y>
- Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., Du, Y., Li, Y., Lee, H., Zheng, H. S., Ghafouri, A., Menegali, M., Huang, Y., Krikun, M., Lepikhin, D., Qin, J., ... Le, Q. (2022). *LaMDA: Language models for dialog applications*. ArXiv Preprint ArXiv:2201.08239. <https://doi.org/10.48550/arXiv.2201.08239>
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *Llama: Open and efficient foundation language models*. ArXiv Preprint ArXiv:2302.13971.
- University of Aberdeen. (n. d.). *Common Grading Scale (CGS)*.

<https://www.abdn.ac.uk/students/academic-life/common-grading-scale.php>

Wang, H., Wu, W., Dou, Z., He, L., & Yang, L. (2023). Performance and exploration of ChatGPT in medical examination, records and education in Chinese: Pave the way for medical AI. *International Journal of Medical Informatics*, 177, 105173. <https://doi.org/10.1016/j.ijmedinf.2023.105173>

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q.-L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica*, 10(5), 1122–1136. <https://doi.org/10.1109/JAS.2023.123618>

Zarb, M., McDermott, R., Martin, K., Young, T., & McGowan, J. (2023). Evaluating a pass/fail grading model in first year undergraduate computing. *2023 IEEE Frontiers in Education Conference (FIE)*, 1–9. <http://dx.doi.org/10.1109/FIE58773.2023.10343276>

## Appendices

### Appendix A

BMS services is a small-scale healthcare company. You have been asked by the team to explore machine learning techniques to analyse and evaluate heart failure model performance to make informed decisions on patient's survival. For this purpose, you will make use of programming and statistical concepts for analysis, visualisation, and machine learning algorithms.

Please download the health failure clinical dataset via this link below  
<https://archive.ics.uci.edu/dataset/519/heart+failure+clinical+records>

Specifically, you are required to perform the following tasks below.

- i). Exploratory Data Analysis  
 You are required to check if there's any missing data. List appropriate data pre-processing steps. Perform required descriptive statistical analysis. Give detailed explanation of all processes.
- ii). Classification I  
 Split the dataset on training and testing sets. You are required to fit all machine learning algorithms namely, Naïve Bayes, Logistic Regression, Support Vector Machine, Random Forest classifier, K-Nearest Neighbour and Multi-Layer Perceptron Neural Networks. Evaluate your models using test dataset and provide the confusion matrix for all models. Report and compare performance of the models in terms of accuracy, precision, recall and F1-Score. Draw conclusions and provide recommendations.
- iii). Classification II  
 Investigate class imbalance problem by producing the plot of the target variable class distribution. If there is presence of class imbalance problem, use at least 2 techniques to balance the class distribution (Algorithm or Sampling technique). This means you will have a balanced dataset. Using the balanced dataset, you are required to build classification models using machine learning algorithms namely, Naïve Bayes, Logistic Regression, Support Vector Machine, Random Forest classifier, K-Nearest Neighbour and Multi-Layer Perceptron Neural Networks. Evaluate your models using test dataset and provide the confusion matrix for all models. Report and compare performance of the models in terms of accuracy, precision, recall and F1-Score. Compare your result with the result of II above. Draw conclusions and provide recommendations. Please provide justification for chosen methods.
- iv). Conclusion  
 Critically review solutions ii). and iii). above. Which model will you recommend?

## Appendix B

### Task Summary:

You are a data scientist at Premium Technologies and recently tasked with developing a ML pipeline for one of your Biomedical clients. You are to perform a text classification of these scientific papers into thyroid, colon, or lung. To simplify the problem, it is sufficient to only consider the first 600 characters of each paper text. Afterwards, evaluate the model's performance. Please download the medical text dataset – cancer doc classification provided via the link below.

Kaggle: <https://www.kaggle.com/datasets/falgunipatel19/biomedical-text-publication-classification>

The dataset consists of the following:

labels: labels describing the different types of cancer; provided in column 'Q'  
 text: free text provides in column 'a' and contains medical papers addressing one of the 3 cancer areas

Requirements

- Please perform the following tasks in a well-documented Jupyter notebook:
1. Load medical text dataset.
  2. Perform appropriate data pre-processing steps.
  3. Implement a text classification approach by using a transformer-model (e.g. BERT model) to convert the text into an embedding space.
  4. Train and execute a classification model and evaluate the model's performance.
  5. Depending on your results, describe further optimisation steps that you would consider in order to improve the performance of your model.

### Additional Information

- Although this is a synthetic dataset, you are to operate under the assumption that this is patient confidential data which cannot be shared with externals
- The objective of this challenge is not to develop the best-performing algorithm, but rather to understand your motivation for a specific approach/solution and to understand your thinking about further improvements.

### Report

Prepare a short report (not exceeding two pages including figures) describing your thoughts on why specific steps/methods were used. You are required to submit your report and the codes / Jupyter notebook.

## Appendix C

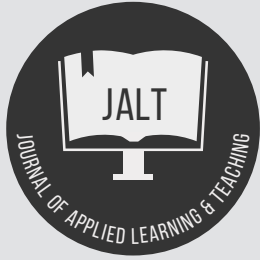
In project performance assessment, quality, cost, and time are the three apexes of the Iron Triangle. Yet, value, safety, and scope are also major criteria. You have been requested to represent your organisation at the 'performance brief' meeting for Quikzome community health center. You are required to create a One-Slide PowerPoint presentation that will illustrate schematically the interrelationship among the project criteria listed above using the Iron Triangle while emphasising the need to avoid or reduce dispute. In addition, give an appropriate title to the schematic illustration.

## Appendix D

Assessment Criteria	Grade A (Excellent)	Grade B (Very Good)	Grade C (Good)	Grade D (Pass)	Grade E (Marginal Fail)	Grade F (Fail)	Grade G (Fail)
Comprehension and Application	Exceptional understanding with nuanced insights. Evidence of extensive reading / research.	Displays a very good understanding of the topic with comprehensive insights. Evidence of reading / research.	Good understanding with some depth and clarity. Evidence of research.	Basic understanding shows with general insights. Evidence of research.	Poor understanding of the topic	Deficiencies in formulation of arguments which are sufficiently serious to indicate a fundamental lack of understanding of the assessment topic.	Token or no submission
Analysis and Synthesis, Accuracy and Relevance	Exceptional analysis of subject with use of advanced theory	Very good analysis of subject with use of relevant theory	Good analysis of subject with use of theory	Some analysis of subject with use of theory	Poor analysis of subject	No appreciable debate/ analysis	Token or no submission
Evaluation/Discussion of Result	Exceptional critical analysis with sophisticated integration of theory and examples.	Provides insightful analysis with well integrated, relevant theory and examples	Good level of analysis with relevant theory and examples.	Some analysis, but mostly descriptive or lacking depth.	Little to no critical analysis.	No appreciable evidence of coherent thinking	Token or no submission
Presentation/Readability	Exceptional clarity, and elegance in writing, error-free	Exceptionally clear, concise, and error-free writing.	Clear writing with minor errors.	Writing is understandable but with several errors.	Poor writing quality with numerous errors.	Consistently poor writing quality/ presentation.	Token or no submission

Copyright: © 2024. Bayode Ogunleye, Kudirat Ibilola Zakariyyah, Oluwaseun Ajao, Olakunle Olayinka and Hemlata Sharma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Awareness, benefits, threats, attitudes, and satisfaction with AI tools among Asian and African higher education staff and students

Muayyad Ahmad <sup>A</sup>	A	<i>Professor, Clinical Nursing Department, School of Nursing University of Jordan</i>
Maha Subih <sup>B</sup>	B	<i>Associate Professor, School of Nursing- Al-Zaytoonah University, Amman-Jordan</i>
Mirna Fawaz <sup>C</sup>	C	<i>Assistant Professor, Nursing Department, Faculty of Health Sciences, Beirut Arab University, Beirut, Lebanon</i>
Hanan Alnuqaidan <sup>D</sup>	D	<i>Assistant Professor, Head of Studies and Information Office, Nursing Department- MOH, State of Kuwait</i>
Ashraf Abuejheisheh <sup>E</sup>	E	<i>Assistant Professor, Department of Nursing, Faculty of Health Profession, Al-Quds University, Jerusalem, Palestine</i>
Vian Naqshbandi <sup>F</sup>	F	<i>Assistant Professor, Adult Nursing College of Nursing/ Hawler Medical University</i>
Fadwa Alhalaiqa <sup>G</sup>	G	<i>Professor, Department of Pre-Clinical Affairs, College of Nursing, QU Health, Qatar University, Doha</i>

### Keywords

Artificial intelligence (AI);  
attitudes;  
awareness;  
ChatGPT;  
higher education;  
intelligence;  
satisfaction;  
threats.

### Abstract

Artificial intelligence (AI) tools are now used in our daily lives. This study aimed to explore the level of awareness, perceived benefits, threats, attitudes, and level of satisfaction with AI tools among individuals within higher education in Asia and Africa. A cross-sectional study was conducted in August 2023. Snowball sampling was used with a convenience sample of 815 highly educated Asian and African participants from 11 countries. About 56% of participants have Bachelor's degrees. 312 participants (38%) were unaware of AI tools and AI tools were used rarely by 316 (63%) of 503 participants who were aware of them. ChatGPT is the most popular of this study's AI tools (N=405, 81%). Participants who used AI tools reported greater benefits than those who did not ( $p < 0.05$ ). Of the four educational groups, those with a Master's degree reported a higher AI tool threat than those with a Diploma ( $P < 0.05$ ). Female participants reported more AI-related threats than males ( $P < 0.05$ ). In conclusion, this research is important because of the rapid development of modern technology around the world. Nevertheless, Asia and Africa still lag behind developed nations in AI technology awareness.

### Correspondence

mma4@ju.edu.jo<sup>A</sup>

### Article Info

Received 20 November 2023  
Received in revised form 7 January 2024  
Accepted 22 January 2024  
Available online 23 January 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.10>

## Introduction

Artificial intelligence (AI) is the technology that enables a computer system or computer-controlled robot to learn, reason, observe, infer, communicate, and make judgments similarly to or better than humans (Crompton & Burke, 2023; Ismail et al., 2023; Robert, 2019). It is one of the most revolutionary technologies of the twenty-first century, with profound effects on the economy and society (Scherer et al., 2023). In recent years, AI has made tremendous advancements, generating a vast array of tools and applications. AI is now an indispensable element of modern civilization. These tools can facilitate decision-making, enhance information transmission, and close knowledge gaps (Rajagopal et al., 2022).

The significance of using AI tools is immense and growing. AI systems improve decision-making by analyzing vast volumes of data to identify patterns and trends that are difficult or impossible for humans to recognize (Aitkazinov, 2023; Bani et al., 2023). These data can be used to improve decisions in various areas, including product development, health care, and customer service. AI technologies that can be tailored to provide each client with the information and services they demand can be used to provide an exceptional customer experience (Adarkwah et al., 2023; Bharadiya, 2023), which could increase client satisfaction (Hu et al., 2023).

Tools for AI are now used in our daily lives (García-Peñalvo, 2023). Education and research based on AI have entered a new, rapidly increasing era. The improvement of research and educational precision and efficiency is one of the primary benefits (Ali et al., 2023; Basilotta-Gómez-Pablos et al., 2022). AI systems can quickly scan large data sets and identify links that are difficult for humans to recognize. This may result in more productive research, allowing scientists to focus on novel and demanding endeavors (Chan & Hu, 2023). For educational reasons, AI systems may assess the learning preferences and aptitudes of students and provide individualized guidance and support to help them achieve their goals (Dergaa et al., 2023).

The evolution of AI has had a major effect on contemporary culture and daily life (Hassoulas et al., 2023). Artificial intelligence has become indispensable, with predictive algorithms improving user experiences and virtual assistants easing activities (Rudolph et al., 2023). It revolutionizes businesses by providing power to recommendation systems, driverless vehicles, and healthcare diagnostics (Chen et al., 2020). However, cautious implementation is required due to ethical considerations, challenges related to data protection, and bias (Bohr & Memarzadeh, 2020). The undeniable impact of AI's ongoing development on our work, communication, and navigation is accompanied by a range of responsibilities and prospects associated with its pervasive integration (Thakur, 2024).

In higher education, AI is revolutionizing the learning environment. By adapting instructional information to the specific requirements of each user, adaptive learning platforms enrich personalized learning experiences (Thakur, 2024). Assessments are streamlined by automated grading systems, which also deliver quick feedback (Hassoulas et al.,

2023). The identification of at-risk students is facilitated by predictive analytics, which permits proactive interventions. Notwithstanding these progressions, ethical deliberations and the imperative for conscientious AI deployment continue to be pivotal in guaranteeing fair and impartial access and upholding the integrity of schooling (Wong et al., 2024).

Furthermore, AI in health studies has grown dramatically in the past decade (Abdullah & Sofyan, 2023). AI can boost healthcare efficiency and affordability. Large IT companies have invested billions in AI research because healthcare uses AI. Technology may replace human interaction and violate care ethics, among other disadvantages (Couture et al., 2023). Technology increases control needs. Healthcare AI has received little legal and ethical scrutiny (Bærøe & Gundersen, 2023).

There may be a limited number of research located in Africa or Asia that are comparable to ours. The benefits, threats, and attitudes of academics regarding AI tools were evaluated using a novel instrument that was constructed based on an extensive literature review. Understanding how faculty members and students use AI tools is essential since their viewpoints and behaviors can have a significant impact on the success of these technologies in their professions. By inquiring about their awareness of AI ideas and comprehension of its potential applications, we may estimate the depth of their expertise in this topic. Thus, this study aims to answer the following questions:

1. What is the level of awareness of AI tools among individuals within higher education in Asia and Africa?
2. Are there differences in the perceived benefits, threats, attitudes, and level of satisfaction with AI tools based on participant characteristics (age, gender, level of education, frequency of use, and country)?

## Methods

### Design

A cross-sectional study was conducted with Asian and African participants, with the majority from the Middle East, during the month of August 2023.

### Sample and setting

Participants represented 11 nations, including nine Asian nations (Jordan, Palestine, Lebanon, Saudi Arabia, Iraq, Thailand, India, Philippines, and Kuwait) and two African nations (Egypt and Sudan). Google Forms were used to collect data from students and faculty at higher education institutions.

In this study, data were collected using a convenience sample because of its practicability and ease of access to participants (Polit & Beck, 2013). The developed online questionnaire link was sent to potential participants in all participating countries via WhatsApp, Facebook Messenger, and email

using the contact lists of the seven researchers of this study who are affiliated with higher education institutions. Each questionnaire link was sent with a message asking them to forward the study questionnaire to their colleagues and students who meet the eligibility requirements.

Eligibility requirements included being a resident of Asia or Africa, possessing a diploma or higher, and/or being a faculty member at a college or university. Because the research instrument was written in English, participants were also required to be able to read English.

## Measures

The study measure consisted of three components. The sociodemographic and personal characteristics are covered in the first section. The second section was designed to measure attitudes, benefits, and threats toward AI tools in higher education. The third section was the satisfaction scale with AI tools in higher education.

The following sociodemographic characteristics were collected: age, gender, country, education, and employment type. Other information pertaining to AI tools was collected such as awareness of different types of AI tools (BardAI, ChatGPT, BingAI, Chatsonic, Writesonic, playground, Claude, Socratic, OpenAI, LaMDA2, Jasper2, and FalconLLM), and frequency of their use.

The psychometric instrument (Appendix 1) was developed by the research team to assess attitudes, benefits, and threats of AI tools among faculty members and students in higher education institutions (Ahmad et al., 2023). The research team reviewed the literature that discussed issues related to AI tools. Each one of the research team extracted main features and then assigned them under the theme of attitudes, benefits, and threats in a draft. The three suggested drafts were merged, and the duplicate items were removed. Then, the psychometric tests were conducted.

Visual Analogue Scale (VAS) was used to determine the level of participant satisfaction with AI tools. The participants were asked to rate their satisfaction with AI tools on a scale from 0 to 100 (Byrom et al., 2022). The higher the score, the more participant satisfaction with AI tools is implied.

## Psychometrics of the attitudes, benefits, and threats instrument

The pre-final draft included 40 items, with seven items to assess the benefits of AI tools, 16 items to assess threats, and 17 items to assess attitude. The research teams agreed on the cleaned version of the instrument. Using the content validity index (CVI), the developed instrument's validity was evaluated. Three experts—one in information technology, one in nursing, and one in medical education—were consulted to determine the validity. The expert panel graded the applicability of each item on the tool. The average of the expert ratings is then used to calculate the CVI. Five of the study's items were removed because their CVI scores were less than .70 or irrelevant. Five experts—three from

the original panel and two new ones from the physics and sociology departments—evaluated the remaining 35 items. The minimum score for each item was .85, and the scale's overall CVI score was .95. Each item was measured using a 5-point Likert scale ranging from strongly disagree (0) to strongly agree (4).

Construct validity assessment on the study scale was performed through exploratory factor analysis (EFA). The 35 items were split up into three factors: 15 items were assigned to attitude, 14 to threat, and 6 to benefits. This analysis's overall explained variance was 55%.

For the three subscales as well as the overall scale, Cronbach's alpha was calculated. The benefits subscale score was 0.82, the threat subscale score was 0.91, and the attitude subscale score was 0.90. Additionally, the overall scale had a reliability score of 0.93.

## Pilot study

A pilot study was conducted during the last week of July 2023, using Google Forms to evaluate the viability of the data collection methods and tools. Thirty faculty members and students were recruited for the piloting using an online survey. The survey comprised sociodemographic and personal data, assessments of attitudes, benefits, and threats regarding AI tools in higher education, as well as a satisfaction scale pertaining to such tools. Both the study scales and the sociodemographic questionnaire were distributed to participants. The principal investigator (PI) observed the participants and recorded any problems with the tools and procedures, such as misinterpretations of the questions or technical problems. The PI also requested participant feedback on the processes and instruments, such as questionnaire length and clarity of instructions.

The data from the pilot study were analyzed to make any necessary adjustments prior to the main study. The minor modifications based on the pilot study included clarification for the first page of the online questionnaire, which served as the consent form's cover sheet, had its font size clarified, the option to select 'other countries' was added, and the option to select the responding student's year of study was removed. The original draft contained 10 common types of AI tools, but pilot participants suggested adding two more, so the final version included twelve types. A final option was added to allow participants who were unfamiliar with the AI tools to submit their responses directly after they complete the sociodemographic questionnaire.

## Ethical considerations

This study was conducted following the ethical guidelines set by the Helsinki declaration (Ashcroft, 2008). The Institutional Review Board (IRB) at the School of Nursing approved the study. The first page of the questionnaire includes the information of the research purpose, method, their rights not to participate and the confidentiality assurance. An email for members of the research team was provided to receive and answer expected participants' questions. The

informed consent was obtained through clicking “yes” for the question “Do you agree to participate in the current study?” The data were saved on the principal investigator’s (PI) desktop, and only authorized research team members had access to the data.

## Data analysis

IBM SPSS 29.01 was used for data analysis (IBM, 2023). Using tables and histograms, descriptive statistics were employed to summarize the demographic and participant characteristics. Analysis of Variance (ANOVA) was used as inferential statistical comparison between the benefits, threats, attitudes, and satisfaction with AI tools and the characteristics of the participants. A post-hoc test was performed on the significant ANOVA results to determine which groups have differences. An Independent sample t-test was used when gender was the independent variable.

## Results

This study included 815 participants with a high level of education from 11 Asian and African countries. Approximately half of the participants (56%) hold a Bachelor’s degree, while nearly 35% hold a Master’s or Doctoral degree. This study has attracted more women than men (57.8%). The age range was from 18 to 69 years. The interesting findings concerned the frequency of AI tool usage and awareness with the 12 most common types of AI tools presented in this study. There were 312 participants (38%) who reported not being aware of any AI tools. In addition, 316 (63%) of the 503 participants who said they were aware of AI tools reported that they had used them rarely (Table 1).

Table 1: Descriptive statistics of the sample (N=815).

Variable	Aware of AI tools N=503		Not aware of any AI tools N=312		All sample N=815	
	n	%	n	%	n	%
<b>Gender</b>						
Male	246	48.9	98	31.4	344	42.2
Female	257	51.1	214	68.6	471	57.8
<b>Age (Mean=30.98, SD=11.71)</b>						
18-20 year	111	22.1	79	25.3	190	23.3
21-30 year	162	32.2	99	31.7	261	32.0
31-40 year	116	23.1	51	16.3	167	20.5
41-69 year	114	22.7	83	26.6	197	24.2
<b>Education</b>						
Diploma	37	7.4	26	8.3	63	7.7
Bachelor	266	52.9	194	62.2	460	56.4
Master	108	21.5	45	14.4	153	18.8
PhD	92	18.3	47	15.1	139	17.1
<b>AI tool usage frequency</b>						
Daily	37	7.4	-	-	38	4.7
Weekly	88	17.5	-	-	90	11.0
Monthly	59	11.7	-	-	59	7.2
Rarely	316	62.8	-	-	316	38.8
Never	-	-	-	-	312	38.3
<b>Country</b>						
Egypt	30	6.0	17	5.4	47	5.8
Sudan	17	3.4	24	7.7	41	5.0
India	17	3.4	29	9.3	46	5.6
Iraq	53	10.5	39	12.5	92	11.3
Jordan	136	27.0	64	20.5	200	24.5
Kuwait	63	12.5	23	7.4	86	10.6
Lebanon	100	19.9	43	13.8	143	17.5
Palestine	67	13.3	45	14.4	112	13.7
Philippine	5	1.0	6	1.9	11	1.3
Saudi Arabia	15	3.0	4	1.3	19	2.3
Thailand	-	-	18	5.8	18	2.2

The 12 different types of AI tools presented in this study’s frequency distribution are shown in Figure 1. ChatGPT appears to be the most well-known and frequently used type (N=405); 81% of the 503 participants who indicated that they are aware of or have used AI tools. Open AI (this is not a type, but the company that owns ChatGPT) (N=173, 34%) was the second most popular mention of AI. About half of the participants (N=255, 51%) claimed to be aware of more than one type of AI tool. Despite being produced by the same company, OpenAI and ChatGPT serve distinct functions (Roumeliotis & Tselikas, 2023). OpenAI Playground is trainable, while ChatGPT is pre-trained and users cannot train it with their own data. While ChatGPT offers a simpler text-based interface for producing natural language responses to user queries, OpenAI Playground gives users a more interactive and visual way to experiment with AI models.

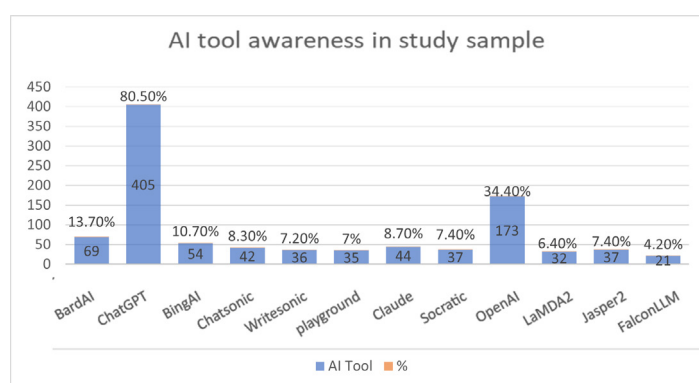


Figure 1: The frequency awareness for the 12 AI tools among the study participants. [What is the meaning of “OpenAI2”?] corrected as OpenAI.

The number of AI tools used, as reported by the 503 participants, is depicted in Figure 2. It is evident that 214 (43%) of the participants have used only one AI tool. Participants who reported using four or more types made up 58 (12%) of the total.

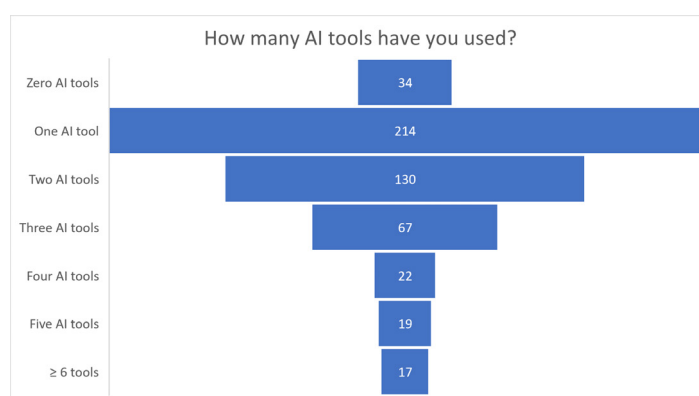


Figure 2: Number of AI tools the participants were aware of and or have used.

Table 2 presents the descriptive statistics for the four study scales. To facilitate interpretation, the original scores for the scales of attitudes, benefits, and threats were transformed to percentages. The final column of the Table shows the range for the four scales of the study between 59.69% and 68.71%. The top two items from each scale are included: for benefits,

AI tools save time and are used in education and research; for threats, AI tools require constant Internet access and reduce creativity and critical thinking; and for attitudes, AI tools may be used if edited and cited appropriately.

Table 2: Descriptive statistics for the study scales and the highest two items in each scale (N=503).

Scales	Mean	Standard deviation	Skewness	Min <sup>^</sup>	Max <sup>^</sup>	% Mean Score
Benefit (6 items)	16.49	4.05	-.82	0	24	68.71
It saves time	2.97	.91				
AI tools are used in education and research	2.80	.98				
Threat (14 items)	36.52	16.9	-.68	0	56	65.24
It needs the Internet all the time	2.87	.99				
It decreases creativity and critical thinking	2.82	1.06				
Attitude (15 items)	38.35	8.87	-.93	0	60	63.92
I review and edit the response that generated by AI tools before using them in my work	2.76	.88				
Can be used if properly cited and documented	2.71	.87				
Satisfaction (VAS)*	-	-	-	0	100	59.69

\*VAS= Visual Analogue Scale, one items %, <sup>^</sup>min=Minimum, <sup>^</sup>max=Maximum

Those who are aware of one or more AI tools reported higher benefits than those who were unaware ( $p < .05$ ). Among the four groups of educational attainment, we found that those with a Master's degree reported a higher AI tool threat than those with a Diploma ( $P < .05$ ). Female participants reported more AI-related threats than males ( $P < .05$ ). There were no significant differences based on education, age groups, gender, or country in terms of the benefits of AI tools. Threats posed by AI tools did not differ significantly by age group, country, or level of awareness. In addition, attitude and satisfaction with AI tools did not differ significantly across all the studied variables (Table 3).

In order to rule out type-II errors in ANOVA analysis and because some countries had small sample sizes, the Kruskal-Wallis test was used to compare the countries against the AI tools in the study. Consistent with the ANOVA analysis, the result indicated that there were no significant differences between the countries.

## Discussion

In the past decade, research and development of AI-based technologies in healthcare, industry, business, and education have increased dramatically. There is a growing awareness of AI tools among faculty and students in higher education institutions around the world (Chan, 2023; Ifelebuegu, 2023). AI tools are becoming more popular for several reasons. Increasing tool availability is one factor. A growing media coverage of AI is another factor. The majority of participants in our study have only used one AI tool, ChatGPT being the most popular. Geographic location affects the degree of familiarity with AI tools. Of the total sample (815 participants), it is interesting to note that 38% of respondents claimed to be unaware of AI tools. Additionally, we found that of the 503 participants who are aware of AI tools, nearly 63% have rarely used it. Compared to the 72% of college students in the United States who are familiar with artificial intelligence (AI) and the 58% who believe AI will have a positive impact on their lives (Rodway & Schepman, 2023). This finding

Table 3: Comparison of the participant characteristics and the benefits, threats, attitudes, and satisfaction with AI tools (N=503).

Characteristics	Benefits of AI tools		Scheffe post hoc
	Statistics test	P value	
Education level	F=0.69	.558	
Age (groups)	F=0.57	.636	
Gender	t=-0.26	.795	
AI tools awareness	F=2.71	.013	All who use AI tools $\geq 1$ , reported higher benefit than those who are not aware about it ( $p < .05$ )
Country	F=0.61	.789	
	Threats of AI tools		
Education level	F=3.92	.009	Master>Diploma ( $p < .05$ )
Age (groups)	F=2.22	.085	
Gender	t=-2.04	.041	Females>Males ( $p < .05$ )
AI tools awareness	F=1.51	.172	
Country	F=1.03	.416	
	Attitudes of AI tools		
Education level	F=1.90	.129	
Age (groups)	F=1.71	.164	
Gender	t=0.30	.768	
AI tools awareness	F=0.69	.657	
Country	F=1.24	.270	
	Satisfaction with AI tools		
Education level	F=1.58	.193	
Age (groups)	F=2.35	.071	
Gender	t=1.71	.089	
AI tools awareness	F=1.39	.339	
Country	F=0.57	.820	

indicates a lack of awareness among our study sample.

According to our findings, 55% of those who used AI tools were under the age of 30, when most people are still enrolled in college as bachelor's or master's students. This could be justified by the fact that people of this age are used to using electronic devices, have grown up in a time where technology is an essential part of daily life, and use these devices for learning, for studying, and even for casual purposes. Many educational institutions have included platforms based on AI into their curricula, exposing students to AI ideas at a young age (Timotheou et al., 2023). Additionally, due to their cognitive flexibility and one of the most economical means of developing young brains, the younger generation is frequently more able to adapt to new technology (Kulkov, 2023).

There are conflicting findings in research regarding gender inequalities in information technology (Liang et al., 2023). In Asia and Africa, men have easier access to IT resources than women do (Kukulka-Hulme et al., 2023). This is consistent with our findings that women made up nearly 69% of participants who were unaware of any AI technologies.

Literature has focused on the possible roles in the medical field, notably in terms of education, research, and clinical settings (Periaysamy et al., 2023). Participants in the current study evaluated AI tools to be beneficial. However, the most troublesome aspect of deploying AI tools is not the level of recognizing AI tools' benefits; rather, the most challenging aspect is proving that AI tools are being used in daily attitude (Himeur et al., 2023).

Academic staff and students in higher education in Asia and Africa have different perspectives on the use of AI tools. Our research found that AI tools pose threats. The main two threats to AI tools were the need for constant Internet access and it also reduced creativity and critical thinking.

The literature, however, recognized many categories of threats. This finding is consistent with previous studies that reported numerous threats posed by AI tools (Alqahtani et al., 2023; Ifelebuegu et al., 2023). In education, students may use AI to write the entire assignment rather than revising it; the data provided by ChatGPT require continuous updating (Periaysamy et al., 2023). Furthermore, Benvenuti et al. (2023) argue that AI tools cannot replace human interaction and that they may lead to a lack of critical thinking skills among students. Furthermore, in the current study, females more than males and those with a master's degree were more likely than those with a diploma degree to perceive AI technologies as threatening. No research has yet directly examined threats in education or by age or gender, but it might be presented as violence due to technology (Novitzky et al., 2023).

Despite the concerns, there are many benefits to using AI tools in higher education. AI can help reduce administrative tasks for teachers and staff, allowing them to focus on more important tasks (Chan, 2023). In our study, the top two benefits for AI tools were saving time and that it is used in education and research. Furthermore, our results support using AI tools if they are edited and cited appropriately. This finding has been emphasized in the recent literature in order to maintain ethics in using AI tools (Atenas et al., 2023; Kooli, 2023).

### Implications

The study could help identify the specific benefits and threats of AI in higher education that are most relevant to students and faculty in Asia and Africa. This knowledge could be used to develop policies and practices that maximize the benefits of AI while minimizing its threats. The research could also assist in determining the level of student and faculty satisfaction with AI in higher education. Furthermore, the study could increase awareness of the potential of AI in higher education; promote the use of AI in higher education in a responsible and ethical manner; inform the development of policies for the use of AI in higher education; and contribute to the body of knowledge on the use of AI in education.

### Conclusion

The application of artificial intelligence in higher education is still in its infancy, but it has the potential to revolutionize how we learn and teach. Asian and African countries included in this study still have lower levels of awareness of AI technology than Asian leaders, like South Korea and China. It is essential to be aware of the potential benefits and threats of AI and to implement safeguards to mitigate the threats. This research is more exhaustive because it includes a large number of participants from eleven Asian and African nations. Investigating 12 common types of AI tools provides valuable insight into the potential benefits, threats, attitudes, and satisfaction with AI tools in education and research. The application of AI in higher education is a difficult and complex issue. However, we must address this issue to ensure that our educational systems are future

ready. Concerns regarding ChatGPT differ based on gender and level of education, despite its increasing global usage. Given the rapid advancement of technology on a global scale, the findings underscore the importance of addressing the lack of awareness regarding artificial intelligence in the studied countries. A limitation of this study could be the non-proportional sampling, despite the fact that the sample size was relatively large and that there was a variety of settings. Moreover, the fact that African respondents originated from two different nations may also limit the generalizability of the findings.

### References

- Abdullah, K. H., & Sofyan, D. (2023). Machine learning in safety and health research: A scientometric analysis. *International Journal of Information Science and Management (IJISM)*, 21(1), 17-37. <https://doi.org/10.22034/ijism.2022.1977763.0>
- Adarkwah, M. A., Amponsah, S., van Wyk, M. M., Huang, R., Tlili, A., Shehata, B., Metwally, A. H. S., & Wang, H. (2023). Awareness and acceptance of ChatGPT as a generative conversational AI for transforming education by Ghanaian academics: A two-phase study. *Journal of Applied Learning and Teaching*, 6(2), 78-93. <https://doi.org/10.37074/jalt.2023.6.2.26>
- Ahmad, M., Alhalaiqa, F., & Subih, M. (2023). Constructing and testing the psychometrics of an instrument to measure the attitudes, benefits, and threats associated with the use of Artificial Intelligence tools in higher education. *Journal of Applied Learning and Teaching*, 6(2), 114-120. <https://doi.org/10.37074/jalt.2023.6.2.36>
- Aitkazhinov, A. (2023). The role of artificial intelligence in auditing: Opportunities and challenges. *International Journal of Research in Engineering, Science and Management*, 6(6), 117-119. <https://doi.org/10.7176/RJFA/14-15-03>
- Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M. A. A., & Dwivedi, Y. K. (2023). A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation & Knowledge*, 8(1), 100333. <https://doi.org/10.1016/j.jik.2023.100333>
- Alqahtani, T., Badreldin, H. A., Alrashed, M., Alshaya, A. I., Alghamdi, S. S., bin Saleh, K., Alowais, S. A., Alshaya, O. A., Rahman, I., & Al Yami, M. S. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy*, 19(8), 1236-1242. <https://doi.org/10.1016/j.sapharm.2023.05.016>
- Ashcroft, R. E. (2008). The declaration of Helsinki. *The Oxford textbook of clinical research ethics*, 141-148.
- Atenas, J., Havemann, L., & Timmermann, C. (2023). Reframing data ethics in research methods education: a pathway to critical data literacy. *International Journal of Educational Technology in Higher Education*, 20(1), 1-27. <https://doi.org/10.1186/s41239-023-00380-y>

- Bærøe, K., & Gundersen, T. (2023). Trustworthy bioethicists within lifecycles of artificial intelligence in health. In *AI and society* (pp. 249-264). Chapman and Hall/CRC.
- Bani Hani, S. H., & Ahmad, M. M. (2023). Machine-learning algorithms for ischemic heart disease prediction: A systematic review. *Current Cardiology Reviews*, 19(1), 87-99. <https://doi.org/10.2174/1573403X18666220609123053>
- Basilotta-Gómez-Pablos, V., Matarranz, M., Casado-Aranda, L.-A., & Otto, A. (2022). Teachers' digital competencies in higher education: a systematic literature review. *International Journal of Educational Technology in Higher Education*, 19(1), 1-16. <https://doi.org/10.1186/s41239-021-00312-8>
- Benvenuti, M., Cangelosi, A., Weinberger, A., Mazzoni, E., Benassi, M., Barbaresi, M., & Orsoni, M. (2023). Artificial intelligence and human behavioral development: A perspective on new skills and competences acquisition for the educational context. *Computers in Human Behavior*, 107903. <https://doi.org/10.1016/j.chb.2023.107903>
- Bharadiya, J. P. (2023). A comparative study of business intelligence and artificial intelligence with big data analytics. *American Journal of Artificial Intelligence*, 7(1), 24. <https://doi.org/10.11648/j.ajai.20230701.14>
- Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. In *Artificial Intelligence in healthcare* (pp. 25-60). Elsevier. <https://doi.org/10.1016%2FB978-0-12-818438-7.00002-2>
- Byrom, B., Elash, C. A., Eremenco, S., Bodart, S., Muehlhausen, W., Platko, J. V., Watson, C., & Howry, C. (2022). Measurement comparability of electronic and paper administration of visual analogue scales: A review of published studies. *Therapeutic Innovation & Regulatory Science*, 56(3), 394-404. <https://doi.org/10.1007/s43441-022-00376-2>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 1-25. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20, 43. <https://doi.org/10.1186/s41239-023-00411-8>
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Couture, V., Roy, M., Dez, E., Laperle, S., & Bélisle-Pipon, J. (2023). Ethical implications of artificial intelligence in population health and the public's role in its governance: perspectives from a citizen and expert panel. *Journal of Medical Internet Research*, 25, e44357. <https://doi.org/10.2196/44357>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), 1-22. <https://doi.org/10.1186/s41239-023-00392-8>
- Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615-622. <https://doi.org/10.5114/biol sport.2023.125623>
- García-Peñalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in the Knowledge Society (EKS)*, 24, e31279-e31279. <https://doi.org/10.14201/eks.31279>
- Hassoulas, A., Powell, N., Roberts, L., Umla-Runge, K., Gray, L., & Coffey, M. (2023). Investigating marker accuracy in differentiating between university scripts written by students and those produced using ChatGPT. *Journal of Applied Learning & Teaching*, 6(2), 71-77. <https://doi.org/10.37074/jalt.2023.6.2.13>
- Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2023). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial Intelligence Review*, 56(6), 4929-5021. <https://doi.org/10.1007/s10462-022-10286-2>
- Hu, K., Lou, B., & Baloch, B. (2023). *Reducing human bias in decision making and improving customer satisfaction through artificial intelligence: Empirical evidence from a knowledge service company*. SSRN 4436554.
- IBM. (2023). *IBM SPSS statistics for Windows, Version 29.0*. Armonk. IBM Corp.
- Ifelebuegu, A. O. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning and Teaching*, 6(2), 385-392. <https://doi.org/10.37074/jalt.2023.6.2.2>
- Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIED) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching*, 6(2), 332-345. <https://doi.org/10.37074/jalt.2023.6.2.29>
- Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, 6(2), 56-63. <https://doi.org/10.37074/jalt.2023.6.2.34>
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7), 5614. <https://doi.org/10.3390/su15075614>
- Kukulka-Hulme, A., Giri, R. A., Dawadi, S., Devkota, K. R., & Gaved, M. (2023). Languages and technologies in education at school and outside of school: Perspectives from young people in low-resource countries in Africa and Asia. *Frontiers in Communication*, 8, 1081155. <https://doi.org/10.3389/fcom.2023.1081155>

Kulkov, I. (2023). Next-generation business models for artificial intelligence start-ups in the healthcare industry. *International Journal of Entrepreneurial Behavior & Research*, 29(4), 860-885. <https://doi.org/10.1108/IJEBR-04-2021-0304>

Liang, M., Lim, C. P., Park, J., & Mendoza, N. B. (2023). A review of ICT-enabled learning for schoolgirls in Asia and its impacts on education equity. *Educational technology research and development*, 71(2), 267-293. <https://doi.org/10.1007/s11423-022-10178-w>

Novitzky, P., Janssen, J., & Kokkeler, B. (2023). A systematic review of ethical challenges and opportunities of addressing domestic violence with AI-technologies and online tools. *Heliyon*, 9(6). <https://doi.org/10.1016/j.heliyon.2023.e17140>

Periaysamy, A. G., Satapathy, P., Neyazi, A., & Padhi, B. K. (2023). ChatGPT: Roles and boundaries of the new artificial intelligence tool in medical education and health research–correspondence. *Annals of Medicine and Surgery*, 85(4), 1317. <https://doi.org/10.1097/MS9.0000000000000371>

Polit, D. F., & Beck, C T. (2013). *Essentials of nursing research: Appraising evidence for nursing practice*. Lippincott Williams & Wilkins.

Rajagopal, N. K., Qureshi, N. I., Durga, S., Ramirez Asis, E. H., Huerta Soto, R. M., Gupta, S. K., & Deepak, S. (2022). Future of business culture: An artificial intelligence-driven digital framework for organization decision-making process. *Complexity*, 2022, 1-14. <https://doi.org/10.1155/2022/7796507>

Robert, N. (2019). How artificial intelligence is changing nursing. *Nursing Management*, 50(9), 30. <https://doi.org/10.1097/01.NUMA.0000578988.56622.21>

Rodway, P., & Schepman, A. (2023). The impact of adopting AI educational technologies on projected course satisfaction in university students. *Computers and Education: Artificial Intelligence*, 100150. <https://doi.org/10.1016/j.caeai.2023.100150>

Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and open-ai models: A preliminary review. *Future Internet*, 15(6), 192. <https://doi.org/10.3390/fi15060192>

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>

Scherer, A. G., Neesham, C., Schoeneborn, D., & Scholz, M. (2023). New challenges to the enlightenment: How

twenty-first-century sociotechnological systems facilitate organized immaturity and how to counteract it. *Business Ethics Quarterly*, 33(3), 409-439. <https://doi.org/10.1017/beq.2023.7>

Thakur, R. (2024). Introduction to artificial intelligence and its importance in modern business management. In *Leveraging AI and emotional intelligence in contemporary business organizations* (pp. 133-165). IGI Global. <https://doi.org/10.4018/979-8-3693-1902-4.ch009>

Timotheou, S., Miliou, O., Dimitriadis, Y., Sobrino, S. V., Giannoutsou, N., Cachia, R., Monés, A. M., & Ioannou, A. (2023). Impacts of digital technologies on education and factors influencing schools' digital capacity and transformation: A literature review. *Education and Information Technologies*, 28(6), 6695-6726. <https://doi.org/10.1007/s10639-022-11431-8>

Wong, B. K. M., Vengusamy, S., & Bastrygina, T. (2024). Healthcare digital transformation through the adoption of artificial intelligence. In *Artificial Intelligence, big data, blockchain and 5G for the digital transformation of the healthcare industry* (pp. 87-110). Elsevier. <https://doi.org/10.1016/B978-0-443-21598-8.00014-2>

## Appendix

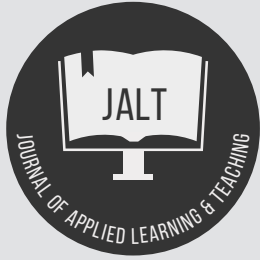
### Appendix 1: Attitudes, benefits, and threats associated with the use of Artificial Intelligence tools in higher education.

Please answer each of the following questions about what you know, how you feel, and what you do with AI tools. (Please note that there is no best answer; we just want to know your opinion about each item.)

	Attitudes (15 items)	Strongly disagree 0	Disagree 1	Neutral 2	Agree 3	Strongly agree 4
A1	AI tools content can be used if properly cited and documented					
A2	Authors should have proper knowledge on how to use AI tools					
A3	I recommend AI tools to a friend or colleague					
A4	I'm interested in using of a premium version of AI tools with advanced features					
A5	AI tools has a positive impact on my education learning					
A6	There is a need for specific training on how to use AI tools in order for them to be useful.					
A7	I suggest providing adequate information on establishing ethical guidelines for the use of AI tools.					
A8	I think AI tools should be included in the study curricula					
A9	To improve AI applications in the real world, it is essential to encourage researchers to be honest and transparent about their methods.					
A10	I review and edit the response that generated by AI tools before using them in my work.					
A11	AI tools can be listed as an author based on its significant contribution					
A12	I feel comfort with ethical and responsible use of AI-generated content from AI tools.					
A13	AI tools could enhance research (e.g., assisting the researchers in framing the sentences, improving the content drafted by the authors.					
A14	I think the responses generated by AI tools are overall easy and coherent					
A15	I trust the information that I read and see on AI tools?					
	<b>Benefits (6 items)</b>	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
B1	Easy to use					
B2	Save time					
B3	Accessible with low cost					
B4	Help students to ask questions and interact with the material at their own pace					
B5	AI tools are user-friendly					
B6	I know that AI tools are used in education and research					
	<b>Threats (14 items)</b>	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
T1	Lack of human interaction					
T2	Legal issue (eg. copyright issues, authorship)					
T3	Decrease creativity and critical thinking					
T4	AI tools does not replace practical training					
T5	Security concerns					
T6	Technical issue					
T7	Over-reliance on technology					
T8	Ethical dilemma concerns such as plagiarism					
T9	Used internet all the time					
T10	Difficulty in handling complex task in research					
T11	Inaccurate/incorrect or biased information					
T12	Over-detailed, redundant, excessive content					
T13	Using AI tools will reduce skills and abilities of person who use it (e.g., writing skills, critical thinking ... etc)					
T14	I see AI tools as a threat to human ethics					

Copyright: © 2024. Muayyad Ahmad, Maha Subih, Mirna Fawaz, Hanan Alnuqaidan, Ashraf Abuejheisheh, Vian Naqshbandi and Fadwa Alhalaiqa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Exploring graduate students' perception and adoption of AI chatbots in Zimbabwe: Balancing pedagogical innovation and development of higher-order cognitive skills

---

Fodouop Kouam Arthur <sup>A</sup> *Hebei University, China*  
William<sup>A</sup>

Muchowe Regis Misheal<sup>B</sup> *Zimbabwe Open University, Zimbabwe*

---

### Keywords

---

Artificial intelligence (AI);  
chatbots;  
graduate students;  
higher-order cognitive skills;  
pedagogical innovation.

---

### Correspondence

willyfodouop@163.com<sup>A</sup>

---

### Article Info

Received 7 December 2023  
Received in revised form 19 January 2024  
Accepted 23 January 2024  
Available online 26 January 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.12>

### Abstract

---

This research paper explores the perceptions and adoption of AI chatbots by graduate students in Zimbabwean universities. The study aims to understand the potential benefits and challenges of adopting AI chatbots in the education sector and its impact on developing higher-order cognitive skills. The study used qualitative methodologies, including field interviews, to analyze the data. The findings suggest that graduate students have a positive attitude towards AI chatbots because they enhance their learning experience, enable them to overcome skill gaps, and aid in test preparation. Furthermore, the study revealed that AI chatbots foster the development of higher-order cognitive skills by augmenting traditional lectures, test preparation, and personalization. However, challenges include plagiarism, outdated information, and financial constraints associated with AI chatbots. The study recommends that AI companies offer discounts to graduate students to enable them to access AI chatbot tools and that universities develop referencing systems allowing students to acknowledge using AI chatbots as sources.

## Introduction

Artificial intelligence (AI) has become widespread in our everyday lives. Preliminary investigations indicate that incorporating AI chatbots into educational settings yields advantageous outcomes for students across multiple domains, including active and constructive learning and creative and social learning (Bii, 2013). Chassignol et al. (2018) believe that AI has a transformative impact on the field of education. Using AI systems and chatbots in education is a potential avenue for advancement (Kooli, 2023). According to Chan and Tsi (2023), integrating AI in educational settings offers valuable enhancements to the teaching and learning processes without being perceived as a substitute for traditional methods. Chatbots are emerging as a novel form of automation, gradually gaining recognition at institutions in Zimbabwe. Graduate programs are vital to Zimbabwe's educational framework; incorporating chatbots to augment the learning experience is a pivotal advancement. As Chang et al. (2023) suggest, incorporating AI chatbots into educational settings has been shown to facilitate and enhance students' self-regulated learning. Pillai et al. (2023) argue that using artificial intelligence and teaching bots in higher education can be attributed to various variables. The authors employ a mixed-method methodology to examine several characteristics, including perceived ease of use, perceived usefulness, interactivity, personalization, perceived intelligence, and anthropomorphism. The integration of chatbots in academic libraries holds significant importance, as Kaushal and Yadav (2022) emphasized, particularly in enhancing research endeavors and facilitating scholarly collaboration. According to Rahim et al. (2022), perceived trust and performance anticipation are the primary factors influencing students' reliance on AI chatbots. In his study conducted in 2021, Malik examines the utilization of chatbots in the context of education. He explores the significance of perceived convenience and improved performance in accepting and integrating chatbots among university students.

However, the higher education industry in Zimbabwe continues to face noteworthy obstacles, including issues related to financial resources and the ability to access educational opportunities. According to Teferra et al. (2013), the issue of funding is a substantial obstacle to the advancement of higher education in Zimbabwe. Furthermore, the rural environment presents barriers to accessing, participating in, and achieving higher education (Nenji & Ndofirepi, 2020). Nevertheless, there has been a notable transition towards incorporating technology within education. According to Isaacs (2007), Zimbabwe has implemented a comprehensive national ICT policy that specifically emphasizes the integration of ICTs within the education sector. Following Tsokota and Solms (2013), it is recommended that the government implement computerization of its processes and establish a favorable climate that enables the private sector to utilize ICT efficiently. The study by Kujeke et al. (2012) examine the difficulties encountered by universities in Zimbabwe when effectively employing ICT for administrative and instructional purposes. The authors highlight the importance of enhancing infrastructure and providing training opportunities for faculty members and students to address these obstacles.

AI chatbots are emerging as a prominent ICT form in Zimbabwean institutions. Lin and Yu (2023) argue that using chatbots in educational settings is a burgeoning area of interest in educational technology research due to its numerous benefits. According to Michel-Villarreal et al. (2023), tailored learning experiences have been found to enhance students' involvement with the learning process. Furthermore, it has been shown that chatbots can offer students prompt feedback, hence enhancing the overall quality of the learning experience (Wang et al., 2023).

Ravankar et al. (2016) underscore the significance of problem-finding skills in teaching problem-solving skills. Furthermore, the study by Lee et al. (2019) examines the extent to which critical thinking and problem-solving abilities are present among undergraduate students in technical fields, emphasizing the necessity for enhancing these skills. Hence, it is imperative to utilize instructional approaches that facilitate the cultivation of cognitive abilities at an advanced level. Incorporating AI chatbots into graduate programs offers a distinct potential to harmonize instructional innovation by cultivating higher-order cognitive abilities.

Implementing and using AI chatbots in Zimbabwe pose hurdles despite their significant potential for aiding the learning process. In their seminal work, Woolf et al. (2013) identify significant issues regarding AI in education. These challenges encompass several aspects, such as the provision of mentors, the acquisition of 21st-century skills, and the establishment of universal access to global classrooms. AI chatbots in education are a nascent development, and the extent to which they enhance the overall quality of instruction and learning outcomes has yet to be definitively established. While there is an increasing body of research in industrialized nations investigating the efficacy of AI chatbots in enhancing educational achievements, there exists a need for more studies examining the impact of AI chatbots on the development of cognitive skills in graduate students.

The research question of this study is formulated as follows: What is the impact of AI chatbots on the development of cognitive skills in graduate students in Zimbabwean universities? This research aims to investigate graduate students' perceptions and attitudes towards using AI chatbots in their learning process. The study also aims to assess the effectiveness of AI chatbots in enhancing the development of cognitive skills among graduate students. Lastly, the potential challenges and limitations associated with using AI chatbots in the educational context will be identified, and recommendations for their integration and effective use will be provided. The rest of this paper is organized as follows. The following section presents an overview of the use of AI chatbots in the education sector. The methodology section outlines the research design, data collection, and data analysis procedures. Section Four highlights the findings and discussions. The last section concludes this study.

## Literature review

Artificial intelligence (AI) technology has rapidly emerged in education, driven by the demand for personalized and adaptive learning experiences. One particular application of AI in education is through the use of chatbots, which have the potential to deliver customized education and support services. This literature review examines the emergence of AI chatbots in education and their impact on student learning processes. It highlights the advantages of using chatbots, such as providing prompt feedback, addressing inquiries, and making tailored suggestions. Moreover, chatbots can assist educators in evaluating assignments, monitoring student progress, and offering administrative assistance. The review also discusses contrasting perspectives regarding the impact of chatbots in education, including concerns about privacy, academic integrity, and the loss of in-person interactions with teachers. While the existing literature showcases the potential benefits of incorporating AI chatbots in education, further research is needed to fully understand their impact on cognitive skills development in graduate students.

### The emergence of AI technology in education

The emergence of AI technology in education can be attributed to the demand for scalable, personalized, and adaptive learning experiences. The integration of chatbots into education has commenced (Winkler & Söllner, 2018). Cunningham-Nelson et al. (2019) assert that chatbots can deliver and customize various aspects of education. According to Okonkwo and Ade-Ibijola (2021), using chatbot technology holds promise in delivering efficient and tailored services to individuals within the educational domain. Furthermore, Verleger and Pembroke (2018) suggested that AI software tools have the potential to revolutionize the student experience. There are several justifications for their acceptance in the educational sector. Sandu and Gide (2019) conducted a study that centers on using AI chatbots inside the higher education system in India. The authors highlight the advantages of utilizing chatbots to improve students' learning experiences. AI chatbots offer students prompt feedback, address inquiries, and make tailored suggestions.

Moreover, chatbots can provide information without requiring extensive, time-consuming searches while concealing the underlying complexity (Ondáš et al., 2019). According to the findings of Pérez et al. (2020), chatbots have the potential to facilitate learning in a manner comparable to that of a human teacher. In addition, Sung (2020) evaluates AI English-language chatbots and posits that they are anticipated to significantly contribute to achieving speaking and listening proficiency benchmarks.

Furthermore, AI chatbots support educators and administrators in their day-to-day responsibilities, particularly evaluating assignments, monitoring student advancement, and offering administrative assistance. According to Yang and Evans (2019), AI chatbots have been found to assist in many educational endeavors. As a result, this technology provides educators with additional time to dedicate to tasks that necessitate human involvement, such as offering mentorship and assistance to students.

Furthermore, the study conducted by Yang (2022) delves into preservice teachers' perspectives regarding AI chatbots in English education. The author's findings shed light on the positive attitudes shown by these individuals towards the efficacy of AI chatbots as valuable tools for teaching and learning purposes. In addition, Thomas (2020) posits that chatbots serve as virtual instructional tools, alleviating instructors from mundane responsibilities. According to the study by Huang et al. (2019), using chatbots in the learning process has been found to mitigate the sense of isolation experienced by e-learners.

Nevertheless, there is a need for improvement in the awareness of AI technology among certain educational institutions. In their study, Adarkwah et al. (2023) examine the level of awareness and degree of acceptance of ChatGPT and AI chatbots among Ghanaian academics. The authors propose that many of the academic community might benefit from acquiring additional knowledge of ChatGPT and chatbots driven by artificial intelligence.

### Discussing the impact of AI chatbots on education and student learning process

The existing scholarly literature offers divergent perspectives on the impact of AI technology on education and students' learning processes. While several scholarly publications showcase the potential advantages of incorporating this technology within education, contrasting research findings indicate a contrary perspective (Gamage et al., 2023; Sullivan et al., 2023).

According to the study conducted by Kim et al. (2021), it is posited that the utilization of AI chatbots has the potential to enhance students' proficiency in English communication within the context of learning English as a Foreign Language. According to Wu and Yu (2023), using AI chatbots has benefited students' learning outcomes, particularly those engaged in higher education. Numerous scholarly publications offer valuable perspectives on the impact of AI chatbots on the educational journey of postgraduate students within higher education institutions. Koivisto (2023) states that implementing chatbots in student counselling can enhance scalability and service hours. However, it is essential to note that students continue to be provided with human counselling services. In their recent publication, Liu et al. (2022) offer a novel chatbot system that utilizes AI to tailor the learning process, augment cognitive capabilities, and boost students' acquisition of learning skills.

In a study conducted by Neo (2022), the MERLIN Project was examined. This project aimed to create a virtual learning assistant utilizing AI chatbot technology. The study results indicated that students evaluated the chatbot as beneficial to their learning experience and effectively enhancing their comprehension of course content. In their study, Pantelić et al. (2023) analyzed student perspectives about AI chatbots and observed a general inclination among students to utilize them for academic-related objectives. According to Chen et al. (2023), using AI chatbots in educational settings facilitates responsive and interactive learning experiences for students, enabling them to acquire crucial material knowledge.

Moreover, these instruments possess significant value in terms of providing educational resources. Hannan and Liu (2023) emphasize AI technology's notable contributions to higher education. As per the authors' assertions, they hold significant importance in facilitating students' learning experiences and providing support to students. According to Michel-Villarreal et al. (2023), incorporating ChatGPT within higher education presents a range of potential benefits for students and educators. These advantages encompass round-the-clock accessibility and assistance, individualized instruction and mentoring, and supplementary educational materials. Additional chances encompass language acquisition and communication proficiency, pedagogical assistance and support for educators and teaching assistants, novel and transformative educational encounters, research endeavors, and data examination.

Furthermore, ChatGPT provides enhanced precision and accuracy in responding to inquiries, generating abstracts, summarizing textual content, and executing various academic-related tasks (Gamage et al., 2023). The usage of chatbots in education is seen favorably by both teachers and students, with the former noting that the chatbot's ability to respond to common inquiries could alleviate their workload (Limna et al., 2023). As per Rasul et al. (2023), ChatGPT can enhance students' learning experiences by helping them develop ideas for their assignments, research, analysis, and assessments. One of ChatGPT's main advantages is that it lets students learn by doing and experiencing things. With ChatGPT, students can assess various methods and techniques for resolving issues and accomplishing objectives through self-directed learning in lieu of a teacher (Rudolph et al., 2023). According to Sullivan et al. (2023), ChatGPT offers distinct possibilities for improving students' academic achievements belonging to diverse equity groups. The study by Yin et al. (2021) investigates the effects of a micro-learning system that utilizes chatbot technology on students' motivation and academic achievement levels. According to the authors, students who incorporate AI chatbots into their educational setting demonstrate high competence and independence. Consequently, these pupils exhibit a reduced need for traditional in-person instruction. Furthermore, these students demonstrate a rapid acquisition of heightened intrinsic desire.

Moreover, Wang et al. (2023) assert that the integration of AI in the realm of higher education affords overseas students the advantage of engaging in individualized and adaptable learning experiences. In addition, artificial intelligence enhances the overall quality of teaching. = Yu (2023) also posits that AI technology presents a significant prospect for education and pedagogy. Specifically, this entails the creation of virtual educational settings and the advancement of virtual instructors. According to Akiba and Fraboni (2023), integrating AI technology is valuable to academic counsellors, fostering educational fairness by empowering individuals individually. Imran and Almusharraf (2023) added that AI chatbots can enhance the efficiency of the academic process.

Nevertheless, chatbots pose potential hazards, including the infringement of privacy and challenges in understanding intricate tasks (Kaushal & Yadav, 2022). Furthermore,

implementing AI technology presents many ethical dilemmas and legal liabilities, most notably academic plagiarism, intellectual property infringement, and the erosion of academic integrity (Yu, 2023). Additional issues of using chatbots in education include the necessity for data security and privacy, the accuracy of the chatbot's information, and the possible loss of in-person interactions with teachers (Limna et al., 2023). According to Rasul et al. (2023), academic dishonesty, prejudice, fabricated data, and poor assessment design will hinder the acquisition of critical graduate skills and encourage cursory learning. According to Perkins (2023), there is a contention about the possible hazards AI Large Language Models pose concerning preserving academic integrity. In a similar vein, Talaue (2023) cautions that the utilization of chatbots by student authors poses a potential threat to the maintenance of academic integrity. Furthermore, the research conducted by Wollny et al. (2021) posits that the evaluation of chatbots with implementation objectives presents significant research issues within the field of education. Michel-Villarreal et al. (2023) employ an ethnographic methodology to shed light on the diverse obstacles associated with using ChatGPT in higher education, particularly emphasizing academic integrity and quality control issues.

Besides, additional concerns warrant attention: personalized learning, expertise, authority, communication, and collaboration. Additionally, integrating AI technology into education has raised significant concerns regarding students' academic performance authenticity. A specific example is the utilization of ChatGPT. It has been identified as a potential avenue for engaging in academic misconduct (Gamage et al., 2023). Similarly, ChatGPT poses a potential risk to academic integrity, particularly concerning plagiarism and academic dishonesty (Sullivan et al., 2023).

The extant body of scholarly work about the impact of AI chatbots on the educational journey of postgraduate students elucidates various advantages and drawbacks associated with their utilization. Nevertheless, further investigation is warranted in the existing scholarly corpus about the equilibrium between educational novelty and the cultivation of advanced cognitive abilities in postgraduate students. Hence, it is imperative to undertake additional research on the impact of AI chatbots on the development of cognitive skills in graduate students.

### **Theory exploring the adoption and use of AI chatbots in education**

The Cognitive Load Theory (CLT) is a theoretical framework explaining how the human mind processes and retains information (Sweller, 1988). It postulates that there are limits to the amount of information the working memory can handle at a given time, and excessive cognitive load can hinder learning and problem-solving. CLT is grounded in the idea that humans have limited cognitive capacities, and the learning process can be optimized by managing and balancing the cognitive load. CLT suggests three types of cognitive load: intrinsic, extraneous, and germane (Sweller, 1988). Intrinsic cognitive load refers to the inherent difficulty of the learning materials or tasks. Extraneous cognitive load

refers to the unnecessary cognitive burden imposed by instructional design elements that do not facilitate learning. Lastly, germane cognitive load refers to the cognitive effort required to construct new schema or knowledge structures.

Previous works relied on CLT to investigate the influence and adoption of AI technology in the educational sector. Abbasi et al. (2019) used the CLT to explore the effects of chatbot systems on students' learning outcomes. In addition, Fidan and Gencel (2022) investigated the impact of chatbots on online learning using the CLT. Similarly, Riapina (2023) analyzed the integration of AI in higher education using the same theory.

This study aims to investigate the impact of germane cognitive load on the potential of AI chatbots to enhance the acquisition of higher-order cognitive skills. AI chatbots can facilitate substantial conversations with pupils, foster the development of critical thinking skills, and assist in tackling intricate problem-solving tasks. AI chatbots can boost germane cognitive load and facilitate the development of higher-order cognitive skills by guiding students during strenuous activities and offering scaffolding support. Hence, this theoretical framework is deemed appropriate for investigating the impact of AI chatbots on the development of cognitive skills in graduate students enrolled in universities in Zimbabwe.

## Methodology and data

The research philosophy of this study is interpretivism because students' perceptions of artificial intelligence and its impact on innovation and higher-order cognitive skills are highly subjective. Therefore, a qualitative methodology was used because the study aimed to explore the influence of artificial intelligence from learners' perspectives. This study's primary data analysis method is thematic analysis (Braun & Clarke, 2006). Thematic analysis identifies and interprets patterns, themes, and meanings within the data. It involves systematically organizing and categorizing data into themes, which helps understand the underlying patterns and relationships. In this research, thematic analysis enables a deep exploration of the participants' perceptions and attitudes toward using AI chatbots. It allows researchers to identify and interpret participants' rich and nuanced responses.

The population is Zimbabwean graduate students in 15 official Zimbabwean universities. A mixture of purposive and snowball sampling was employed. Purposive sampling was adopted because the study targeted only graduate students' views in Zimbabwean universities. On the other hand, the interviews were the ones that referred the researcher to other potential participants who were also university students (snowball sampling) (Reeves et al. 2013). Each interview lasted for about 30 minutes. Because the study is qualitative, saturation point sampling was used, and saturation was achieved at the 15th interview. Therefore, interviews were immediately stopped to save time and resources associated with the research. Critical incident techniques were used during participant interviews to draw on participants' experiences and observations on artificial

intelligence (Tuffour, 2017). Table 1 below displays the profile of the participants.

Table 1. Profile of the participants.

Attribute	Frequency	Percentage (%)
<b>Gender</b>		
Male	7	46.67
Female	8	53.33
<b>Age</b>		
25-35 years	10	66.67
Above 35 years	5	33.33
<b>Employment Status</b>		
Employed	6	54
Unemployed	9	46

Eight participants were women, and seven were men. 10 were aged between 25 and 35 years, and only five were above 35 years old. This is because graduate studies such as Master's and Ph.D. studies are done by older students as compared to undergraduate studies. Only six graduate students were employed at the time of the study. What is essential is that the study was inclusive, and the beliefs and perceptions of diverse groups were considered.

The Guba model was used for data trustworthiness as propounded by Guba and Lincoln in 1994. The model states that qualitative research is subjective and susceptible to participant biases. In addition, to avoid these biases, repeat interviews were done to measure data consistency after two weeks. The first interview was held on the 10th of November 2023, and the second on the 24th of November 2023. The consistent responses made this study credible, dependable, and transferable. NVivo software v14.23.0 was used to organize and analyze qualitative data obtained from field interviews. Data analysis was two-phased, and it started with open coding identifying repeated words as themes and ended with selective coding where themes were solidified and combined depending on relatedness. Ethics were not an issue as participants were kept anonymous and treated as autonomous agents who could withdraw from the study whenever they wished.

## Results and discussions

### Perceptions and attitudes of graduate students towards the use of AI chatbots in their learning process

The study aimed to understand graduate students' perceptions and attitudes toward using AI chatbots in their learning process. The study found that graduate students are aware of AI chatbots and use them in their learning process. One of the participants said this:

My first encounter with chatbots was with FoondaMate, which I used to assist my child in Form Four. However, FoondaMate is for secondary and primary school only. In the class WhatsApp group, a friend shared a WhatsApp chatbot contact (+27600703213), PI [Personal Intelligence chatbot], covering all primary,

secondary, and tertiary education. I have been using it for my assignment tasks and research. (R4)

However, graduate students are more averse to chatbots on social media (WhatsApp and Telegram). They needed to ponder their awareness of ChatGPT, a more advanced AI. Few of them have been using ChatGPT, and they gave two reasons: ChatGPT was not available in Zimbabwe and is more expensive given the economic hardships in Zimbabwe. Below is an extract from one of the students who summarized why graduate students in Zimbabwe tend to use something other than ChatGPT.

ChatGPT is unpopular in Zimbabwe because, since its inception last month, one could not open a ChatGPT account. Only those who use VPNs and have foreign phone numbers can use ChatGPT. However, this is a hustle; many students do not open the accounts. However, now one can open ChatGPT, but you are limited to GPT-3, which is [the free version]. To access GPT-4, which is advanced, one has to pay a subscription of US\$20 per month, and this is beyond the reach of many graduate students, given our economy. (R7)

In a nutshell, Zimbabwean graduate students are using AI in their learning processes (assignments and research). However, they are using chatbots that are free and available on WhatsApp and Telegram. ChatGPT is not used in Zimbabwe because up until last month, Zimbabweans could not open accounts, and now that it is there, the subscription fee for GPT-4 is expensive to them. This finding deviates from Yu (2023) and Wang et al. (2023), who found that ChatGPT is the preferred AI platform for learning. This deviation may be because Yu (2023) and Wang et al. (2023) conducted their studies in developed countries.

In terms of perceptions and attitudes, the study found that most graduate students have a positive perception and attitude towards AI. The positive attitude is because of AI's positive impact on their learning process. For example, R2 is happy about AI because it is "a source of research".

I am studying data analytics, and in Zimbabwe, this is a new emerging field, and there are limited textbooks both physical and in online libraries... Before using chatbots, I struggled to find information about my study. If I want to know a certain concept, I type on a WhatsApp chatbot, which can tell me the concept using simple terms. (R2)

On the other hand, R5 is enthusiastic about AI because it has brought about efficiency in completing tasks such as assignment writing.

AI chatbots help me in writing my assignments. Usually, my assignments are in essay format. The chatbot writes the whole essay; my role is to put Zimbabwean examples only, which these chatbots are limited in. An assignment that took me one month to write now takes me five minutes. (R5)

R5 was not worried about hallucinations of chatbots and ethics involved in using AI to write the whole assignment. R5 maintained that using AI is more ethical than copying and pasting from the internet. R5 said that copying and pasting from another human is undermining the hard work of others while copying and pasting from AI is not the same because no humans are involved. Some graduate students are happy because AI helps in paraphrasing. R7 is of the view that AI solves all plagiarism dilemmas.

I used to struggle with reducing the similarity index in my assignments and research projects. I now use Quillbot to paraphrase... so that my similarity index complies with the 10% accepted rate. (R7)

Some graduate students positively perceive AI because it helps them resolve grammatical challenges. Learning is done in English in Zimbabwe, and English is the second language of many students, and there are bound to be grammatical errors. R9 clarified this.

What frustrates our professors is marking assignments with grammatical issues. The grammar used to result in me getting low marks. These days, before I submit my assignment, I upload it into the chatbot for editing, and recently, my assignment marks have improved. (R9)

The study concludes that many graduate students have a positive perception and attitude towards AI because it has brought about efficiency, helps them address grammatical errors, is a source of research, and helps in paraphrasing. This finding converges with findings by Michel-Villarreal et al. (2023) in Spain and Hannan & Liu (2023) in China that students are bound to have positive attitudes and perceptions of AI because it makes learning easier particularly for graduate students who have other commitments, such as family and work.

### **Effectiveness of AI chatbots in fostering the development of higher-order cognitive skills among graduate students**

The final objective of the study is to understand the effectiveness of AI chatbots in fostering higher-order cognitive skills among graduates. In terms of this objective, three themes emerged: the augmentation of traditional lectures, test preparation, and personalization.

#### ***Augmentation of traditional lectures***

The investigation found that AI chatbots are effective as they augment traditional lectures. Graduate students stated that they use AI chatbots to understand concepts they missed out during traditional lectures.

Before a lecture, I prepare myself for the topics we will learn using chatbots... This has made me participative during lectures, and my lecturers have commented that I have improved in my classroom participation. (R2)

However, R15, unlike R2, uses chatbots after lectures to seek further clarity and understanding.

I am an introvert, and I miss out on much information during lectures because I am too shy to ask the lecturer for further clarity... After the lesson, I then asked the chatbots those questions to get further clarity. (R15)

This finding is different from Chen et al. (2023), who found that AI will replace traditional learning. This study observed that AI chatbots and traditional learning can complement learners' cognitive development.

### **Test preparation**

Some graduate students highlighted that chatbots are pivotal to cognitive development because they aid in examination preparation. It was highlighted by R2, who was enthusiastic about this.

The most preferred revision method was group discussions, but I could not attend group discussions because of family commitments and the costs associated with travelling to the discussion venue. I now use chatbots, and I can write a mock test, and the chatbot can mark it for me. I do this until I feel ready for the examination and the chatbot gives me the corrections. (R2)

This finding corresponds to Wu and Yu (2023), who argue that AI is critical in learners' cognitive development because it helps them prepare for examinations, which is an essential part of the cognitive development of graduate students.

### **Personalization**

One other area that was unearthed by the research is that AI chatbots are vital in cognitive development. Graduate students in the research explained that in AI chatbots, you can ask them what you want in your own time, which is vital in cognitive development. R11, concerned about the content, highlighted this:

The thing with AI chatbots is you can ask them anything that you want, unlike in a lecture where someone asks the lecturer about concepts that you already know, and we end up not having time to ask the lecturer what we do not know. Chatbots allow us to ask about concepts we do not know and get enhanced understanding. (R11)

However, R13 is happy about personalization in terms of time. With lectures, there is a time limit. For example, lecturers prefer students to book appointments to attend to them; this is not the case with AI chatbots.

Chatbots are very flexible whenever I am studying; it can be in the middle of the night, I can ask the chatbot about a concept I do not understand, and I always get feedback. (R13)

In summary, AI chatbots give graduate students the ability to personalize their cognitive development, a concept not identified, for instance, by studies conducted by Liu et al. (2022) and Wang et al. (2023), making this finding an under-researched area in AI chatbot studies and pedagogical innovation.

However, some graduate students noted some challenges associated with AI chatbots, making them less effective in their academic cognitive development. For example, R10 spoke about how AI has resulted in plagiarism.

Some students are no longer doing their assignments; they task the chatbot and copy and paste. It is making graduate students lazy, and they now lack essential cognitive skills such as problem-solving, which are essential. (R10)

It is in line with studies by Hannan and Liu (2023) that show that with AI, there is always a temptation for students to plagiarize. Students are becoming lazy with the introduction of AI chatbots in education (Michel-Villarreal et al., 2023).

However, R6 believes that AI chatbots are ineffective in cognitive development because they are too general. ChatGPT-3.5 does not have up-to-date information on Zimbabwean issues, and it is also prone to hallucinations.

The chatbot I use has limited information. If you want to learn about a recent phenomenon, it tells you that it only has information up to 2021, and this is a challenge; we have to go back to our lecturers for information.

However, this deviates from Wu and Yu (2023), who found that AI chatbots have up-to-date information essential for students. The reason is most likely because they use GPT-4, which has access to the Internet, especially when plugins are used. This deviation may be because of the free AI chatbots (like ChatGPT-3.5) that tend to be used by graduate students in Zimbabwe.

Many graduate students who participated in the study asserted that GPT-4 is expensive, as graduate students cannot afford to pay US\$20 per month, given the harsh economic environment in Zimbabwe. However, this deviates from Yu (2023), who found that the current US\$20 more affordable than the initial US\$40. This divergence is due to the differences in economic contexts.

### **Conclusions and recommendations**

This research paper aimed to explore the influence of AI chatbots on the learning process of graduate students in Zimbabwean universities and their impact on balancing pedagogical innovation with the development of higher-order cognitive skills. Through qualitative methodologies, including field interviews, the study aimed to better understand graduate students' perceptions and attitudes towards AI chatbots and their effectiveness in fostering cognitive skills.

The study found that graduate students know and use AI chatbots in their learning process. However, they are mostly limited to free chatbots available on social media platforms such as WhatsApp and Telegram. The study also uncovered a positive perception and attitude towards AI chatbots, as students appreciate their role in making research more accessible, enhancing assignment efficiency, improving grammar, and enabling personalization in their learning. The findings suggest that AI chatbots foster higher-order cognitive skills by augmenting traditional lectures, test preparation, and personalization.

However, the study identified challenges associated with integrating AI chatbots in education, including plagiarism, outdated information, and financial constraints. It is recommended that AI companies provide discounts to graduate students to enable them to access advanced AI chatbot tools, and universities develop academic policies that allow students to acknowledge using AI chatbots as sources. There is a need to democratize access to AI chatbots by making them affordable and available to students in Zimbabwean universities. It can be achieved through partnerships between AI companies, universities, and the government to ensure that AI chatbots are accessible to all students, regardless of their socio-economic backgrounds. In addition, universities should provide training opportunities for lecturers to enable them to integrate AI chatbots effectively into their teaching methods and facilitate the development of higher-order cognitive skills in graduate students.

Furthermore, universities in Zimbabwe should integrate AI chatbots into formal courses to enhance the learning experience and foster the development of higher-order cognitive skills. This integration should not replace traditional teaching methods but should be used to augment classroom teachings. Lastly, there is a need to develop ethical guidelines for using AI chatbots in education, specifically regarding academic integrity and plagiarism.

Despite this research's valuable insights, several limitations should be acknowledged, notably the focus on graduate students, the limited geographical scope, and the lack of comparative analysis. The study could be expanded further to include undergraduate students and their experiences with AI chatbots to gain a more comprehensive understanding of the impact of these tools on overall education. Moreover, future research should consider conducting cross-cultural studies to explore the impact of AI chatbots on educational practices in different contexts. Lastly, further research could explore the effectiveness of AI chatbots compared to traditional teaching methods or other technological tools.

## References

- Abbasi, S., Kazi, H., & Hussaini, N. N. (2019). Effect of chatbot systems on student's learning outcomes. *Sylwan*, 163(10), 49-63.
- Adarkwah, M. A., Amponsah, S., van Wyk, M. M., Huang, R., Tlili, A., Shehata, B., Metwally, A. H. S., & Wang, H. (2023). Awareness and acceptance of ChatGPT as a generative conversational AI for transforming education by Ghanaian academics: A two-phase study. *Journal of Applied Learning and Teaching*, 6(2), 78-93. <https://doi.org/10.37074/jalt.2023.6.2.26>
- Akiba, D., & Fraboni, M. C. (2023). AI-supported academic advising: Exploring ChatGPT's current state and future potential toward student empowerment. *Education Sciences*, 13(9), 885. <https://doi.org/10.3390/educsci13090885>
- Bii, P. K. (2013). Chatbot technology: A possible means of unlocking student potential to learn how to learn. *Educational Research*, 4(2), 218-221.
- Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Chan, C. K. Y., & Tsi, L. H. (2023). *The AI revolution in education: Will AI replace or assist teachers in higher education?* arXiv:2305.01185. <https://doi.org/10.48550/arXiv.2305.01185>
- Chang, D. H., Lin, M. P. C., Hajian, S., & Wang, Q. Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability*, 15(17), 12921. <https://doi.org/10.3390/su151712921>
- Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science*, 136, 16-24. <https://doi.org/10.1016/j.procs.2018.08.233>
- Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, 25(1), 161-182. <https://doi.org/10.1007/s10796-022-10291-4>
- Cunningham-Nelson, S., Boles, W., Trouton, L., & Margerison, E. (2019). A review of chatbots in education: Practical steps forward. *30th Annual conference for the Australasian Association for Engineering Education (AAEE 2019): Educators becoming agents of change: Innovate, integrate, motivate* (pp. 299-306). Engineers Australia.
- Fidan, M., & Gencel, N. (2022). Supporting the instructional videos with chatbot and peer feedback mechanisms in online learning: The effects on learning performance and intrinsic motivation. *Journal of Educational Computing Research*, 60(7), 1716-1741. <https://doi.org/10.1177/07356331221077901>
- Gamage, K. A., Dehideniya, S. C., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: More opportunities than concerns? *Journal of Applied Learning and Teaching*, 6(2), 358-369. <https://doi.org/10.37074/jalt.2023.6.2.32>
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of Qualitative Research*, 2(163-194), 105.



- Hannan, E., & Liu, S. (2023). AI: New source of competitiveness in higher education. *Competitiveness Review: An International Business Journal*, 33(2), 265-279. <https://doi.org/10.1108/CR-03-2021-0045>
- Huang, W., Hew, K. F., & Gonda, D. E. (2019). Designing and evaluating three chatbot-enhanced activities for a flipped graduate course. *International Journal of Mechanical Engineering and Robotics Research*. <https://doi.org/10.18178/ijmerr.8.5.813-818>
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), 1-14. <https://doi.org/10.30935/cedtech/13605>
- Isaacs, S. (2007). *Survey of ICT and education in Africa: Zimbabwe country report*. InfoDev ICT and Education Series Washington, World Bank Group. <http://documents.worldbank.org/curated/en/377851468007810842/Survey-of-ICT-and-education-in-Africa-Zimbabwe-country-report>
- Kaushal, V., & Yadav, R. (2022). The role of chatbots in academic libraries: An experience-based perspective. *Journal of the Australian Library and Information Association*, 71(3), 215-232. <https://doi.org/10.1080/24750158.2022.2106403>
- Kim, H. S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korea Journal of English Language and Linguistics*, 21, 712-734. <https://doi.org/10.15738/kjell.21..202108.712>
- Koivisto, M. (2023). Tutoring postgraduate students with an AI-based chatbot. *International Journal of Advanced Corporate Learning*, 16(1), 41-54. <https://doi.org/10.3991/ijac.v16i1.35437>
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7), 5614. <https://doi.org/10.3390/su15075614>
- Kujeke, E. C., Thomas, K. A., & Nyaruwata, L. (2012). An impact evaluation of ICT utilization in Zimbabwean universities' ODL administrative and delivery processes. A case of 4 Harare universities (2009-2011). *International Journal of Science and Research (IJSR)*, 3(6), 2267-2275.
- Lee, M. F., Sohod, S. N. M., & Ab Rahman, A. (2019). Exploring the mastery level of critical thinking and problem-solving skill among the technical undergraduate. *Journal of Technical Education and Training*, 11(3). <https://doi.org/10.30880/jtet.2019.11.03.002>
- Limna, P., Kraiwani, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching*, 6(1), 64-74. <https://doi.org/10.37074/jalt.2023.6.1.32>
- Lin, Y., & Yu, Z. (2023). A bibliometric analysis of artificial intelligence chatbots in educational contexts. *Interactive Technology and Smart Education*. <https://doi.org/10.1108/ITSE-12-2022-0165>
- Liu, L., Subbareddy, R., & Raghavendra, C. G. (2022). AI intelligence chatbot to improve students learning in the higher education platform. *Journal of Interconnection Networks*, 22(Supp02). <https://doi.org/10.1142/S0219265921430325>
- Neo, M. (2022). The Merlin project: Malaysian students' acceptance of an AI chatbot in their learning process. *Turkish Online Journal of Distance Education*, 23(3), 31-48. <https://doi.org/10.17718/tojde.1137122>
- Malik, R., Shrama, A., Trivedi, S., & Mishra, R. (2021). Adoption of chatbots for learning among university students: Role of perceived convenience and enhanced performance. *International Journal of Emerging Technologies in Learning (IJET)*, 16(18), 200-212. <https://doi.org/10.3991/ijet.v16i18.24315>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>
- Nenji, S., & Ndofirepi, A. P. (2020). Rurality in higher education in Zimbabwe: Access, participation and achievement. In A. P. Ndofirepi & A. Masinire (Eds.), *Rurality, social justice and education in Sub-Saharan Africa volume II*. Palgrave Macmillan, Cham. [http://dx.doi.org/10.1007/978-3-030-57215-0\\_1](http://dx.doi.org/10.1007/978-3-030-57215-0_1)
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Ondáš, S., Pleva, M., & Hládek, D. (2019, November). How chatbots can be involved in the education process. In *2019 17th International Conference on Emerging Elearning Technologies and Applications (ICETA)* (pp. 575-580). IEEE. <http://dx.doi.org/10.1109/ICETA48886.2019.9040095>
- Pantelić, N., Milošević, M., & Marković, V. B. (2023). Using AI chatbots in academia - The opinions of university students. *SINTEZA*. <http://dx.doi.org/10.15308/Sinteza-2023-306-311>
- Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28(6), 1549-1565. <http://dx.doi.org/10.1002/cae.22326>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2), 07. <http://dx.doi.org/10.53761/1.20.02.07>
- Pillai, R., Sivathanu, B., Metri, B., & Kaushik, N. (2023). Students' adoption of AI-based teacher-bots (T-bots) for learning in higher education. *Information Technology & People*, 37(4). <http://dx.doi.org/10.1108/ITP-02-2021-0152>

- Rahim, N. I. M., Iahad, N. A., Yusof, A. F., & Al-Sharafi, M. A. (2022). AI-based chatbots adoption model for higher-education institutions: A hybrid PLS-SEM-Neural network modelling approach. *Sustainability*, 14(19), 12726. <http://dx.doi.org/10.3390/su141912726>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., ... & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1), 41-56. <http://dx.doi.org/10.37074/jalt.2023.6.1.29>
- Ravankar, A. A., Imai, S., Shimamura, M., Chiba, G., Takasuka, T., & Yamanaka, Y. (2016, July). Nurturing problem-finding skills in graduate students through problem-based learning approaches. In *2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)* (pp. 542-546). IEEE. <http://dx.doi.org/10.1109/IIAI-AAI.2016.177>
- Reeves, S., Peller, J., Goldman, J., & Kitto, S. (2013). Ethnography in qualitative educational research. *Medical Teacher*, 35(8), 1365-1379. <http://dx.doi.org/10.3109/0142159X.2013.804977>
- Riapina, N. (2023). Teaching AI-enabled business communication in higher education: A practical framework. *Business and Professional Communication Quarterly*, 23294906231199249. <http://dx.doi.org/10.1177/23294906231199249>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sandu, N., & Gide, E. (2019, September). Adoption of AI-chatbots to enhance student learning experience in higher education in India. In *2019 18th International conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1-5). IEEE.
- Sullivan, M., Kelly, A., & McLaughlin, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31-40. <http://dx.doi.org/10.37074/jalt.2023.6.1.17>
- Sung, M. C. (2020). Pre-service primary english teachers' AI chatbots. *Language Research*, 56(1), 97-115. <http://dx.doi.org/10.30961/lr.2020.56.1.97>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. [https://doi.org/10.1016/0364-0213\(88\)90023-7](https://doi.org/10.1016/0364-0213(88)90023-7)
- Talaue, F. G. (2023). Dissonance in generative AI use among student writers: How should curriculum managers respond? In *The 5th International Conference of Biospheric Harmony Advanced Research (ICOBAR 2023)* (Vol. 426, p. 01058). EDP Sciences. <http://dx.doi.org/10.1051/e3sconf/202342601058>
- Teferra, D., Mpofu, J., Chimhenga, S., & Mafa, O. (2013). Funding higher education in Zimbabwe: Experience, challenges and opportunities of the cadetship scheme. In D. Teferra (Eds.), *Funding higher education in sub-Saharan Africa* (pp. 327-350). Palgrave Macmillan. [http://dx.doi.org/10.1057/9781137345783\\_13](http://dx.doi.org/10.1057/9781137345783_13)
- Thomas, H. (2020). Critical literature review on chatbots in education. *International Journal of Trend in Scientific Research and Development*, 4(6), 786-788.
- Tsokota, T., & Solms, R. (2013). ICT and the turning-around of the Zimbabwean Economy. In *International conference on ICT for Africa* (pp. 20-23).
- Tuffour, I. (2017). A critical overview of interpretative phenomenological analysis. *A Contemporary Qualitative Research Approach*, 2(4), 52(1) -52(5). <http://dx.doi.org/10.4172/2472-1654.100093>
- Verleger, M., & Pembridge, J. (2018). A pilot study integrating an AI-driven chatbot in an introductory programming course. In *2018 IEEE Frontiers in Education Conference (FIE)* (pp. 1-4). IEEE. <http://dx.doi.org/10.1109/FIE.2018.8659282>
- Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of Artificial Intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences*, 13(11), 6716. <http://dx.doi.org/10.3390/app13116716>
- Winkler, R., & Söllner, M. (2018, July). Unleashing the potential of chatbots in education: A state-of-the-art analysis. In *Academy of management proceedings*, (vol. 2018, no. 1, p. 15903). Briarcliff Manor. <http://dx.doi.org/10.5465/AMBPP.2018.15903abstract>
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachler, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence*, 4, 654924. <http://dx.doi.org/10.3389/frai.2021.654924>
- Woolf, B. P., Lane, H. C., Chaudhri, V. K., & Kolodner, J. L. (2013). AI grand challenges for education. *AI Magazine*, 34(4), 66-84. <http://dx.doi.org/10.1609/aimag.v34i4.2490>
- Wu, R., & Yu, Z. (2023). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13334>
- Yang, J. (2022). Perceptions of preservice teachers on AI chatbots in English education. *International Journal of Internet, Broadcasting and Communication*, 14(1), 44-52.
- Yang, S., & Evans, C. (2019, November). Opportunities and challenges in using AI chatbots in higher education. In *Proceedings of the 2019 3rd international conference on education and e-learning* (pp. 79-83). <https://doi.org/10.1145/3371647.3371659>
- Yin, J., Goh, T. T., Yang, B., & Xiaobin, Y. (2021). Conversation technology with micro-learning: The impact of chatbot-based learning on students' learning motivation and

performance. *Journal of Educational Computing Research*, 59(1), 154-177. <https://doi.org/10.1177/0735633120952067>

Yu, H. (2023). Reflection on whether ChatGPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*, 14, 1181712. <https://doi.org/10.3389/fpsyg.2023.1181712>

Copyright: © 2024. Fodouop Kouam Arthur William and Muchowe Regis Misheal. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



---

**Artificial Intelligence (AI) in academic research. A multi-group analysis of students' awareness and perceptions using gender and programme type**

---

Ofem Usani Joseph <sup>A</sup>	A	<i>Department of Educational Foundations, University of Calabar, Nigeria</i>
Iyam Mary Arikpo <sup>B</sup>	B	<i>Department of Vocational Education, University of Calabar, Nigeria</i>
Ovat Sylvia Victor <sup>C</sup>	C	<i>Department of Educational Foundations, University of Calabar, Nigeria</i>
Nworgwugwu Esther Chidirim <sup>D</sup>	D	<i>Department of Educational Foundations, University of Calabar, Nigeria</i>
Anake Paulina Mbua <sup>E</sup>	E	<i>Department of Guidance and Counselling, University of Calabar Calabar, Nigeria</i>
Udeh Maryrose Ify <sup>F</sup>	F	<i>Department of Science Education, University of Calabar, Calabar, Nigeria</i>
Otu Bernard Diwa <sup>G</sup>	G	<i>Department of Educational Foundations, University of Calabar, Nigeria</i>

---

**Keywords**

Artificial intelligence;  
awareness;  
multigroup analysis;  
perception;  
structural equation modelling;  
utilisation.

---

**Abstract**

The era of AI has brought tremendous impact in academic research, and this has provided the impetus for students to leverage on novel tools in carrying out a lot of quality research works. Previous studies have relied so much on AI for instruction, classroom management and assessment and utilisation of AI tools for research has scarcely been examined. This study covered the gap by examining students' utilisation of AI tools based on their level of awareness and perception and finding out the difference based on gender and programme type in such prediction. A total of 5554 university students were used for the study. Exploratory factor analysis was first carried for dimensionality and other validity checks (convergent and discriminant) using Average Variance Extracted (AVE) and Fornel-Larcker criterion and methods. Population t-tests and multi-group analyses were performed using SPSS and Smart PLS 3. The study found that students have high level of awareness and positive perception of AI tools in research. Similarly, the level of utilisation of AI tools in research is high. Male and postgraduate students have a higher level of awareness and positive perception of AI tools in research, with female students stronger than male students in terms. Perception and awareness directly impacted on utilisation but perception mediates positively and significantly in the nexus between awareness and utilisation. The study findings provide useful insights into using AI tools among university students and also identify the rationale to consider variables like gender and programme type when developing curriculum that will meet the current technology needs in our higher institutions.

---

**Correspondence**

ofemoracle@gmail.com<sup>A</sup>

---

**Article Info**

Received 7 December 2023  
Received in revised form 20 December 2023  
Accepted 4 January 2024  
Available online 22 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.9>

## Introduction

Technology in research has been a product of recurrent invention, but in recent years, AI, which performs cognitive tasks that are problem-solving-oriented, has been commendable (Bonk & Wiley, 2020). The term AI is a conglomeration of different analytical methods classified as machine learning, neural networks, and deep learning (Alloghani et al., 2020; Popenici & Kerr, 2017; Aggarwal, 2018). Each of these concepts has a function it plays. For example, machine learning is programmed with the internal capacity to make decisions through supervised or unsupervised learning models.

The benefits of AI in educational circles and research have been well documented in previous studies. For example, it is stated that AI technology like chatbots is used for review of literature (Clark, 2020), intelligent tutoring and automated data collection (Heffernan & Heffernan, 2014), student collaboration and personalisation of learning experiences (Luckin et al., 2016; Chiu et al., 2022; Mertala et al., 2022), monitoring progress of a work (Swiecki et al., 2019), automated data collection and analysis (Okada et al., 2019; Vij et al., 2020; Yuan et al., 2020), profiling respondents' background (Cohen et al., 2017), as well as analysis of data using different statistical packages (Owan et al., 2023). Other areas are AI tools like ChatGPT that have the internal capacity to assist students and researchers, in areas which include writing tasks, text generation, language translations, and responding to academic queries (Dwivedi et al., 2023; Kasneci et al., 2023; Lund et al., 2023). Similarly, AI utilisation in academic research is effective in assisting students to review literature, overcome barriers in English, usually from those of a non-English speaking background, summarise papers, identify gaps for reviews, and generate drafts of research papers (Rahman et al., 2023; Gao et al., 2022; Rudolph et al., 2023a).

Until recent, many scholars have agreed that AI tools are powerful in improving students writing skills (Zhao, 2022; Kurniati & Fithriani, 2022; Wang, 2022) while others have noted that even though it is very important in improving students' skills, its side effect is considered very paramount (Lund & Wang, 2023; Qadir, 2022; Liu et al., 2022). Given the relevance of AI in research, one anticipates that students' utilisation of this technology in their academic research activities will be maximised. Similarly, their perception will be positive. However, the utilisation of AI is below expectations among students, unlike what it is used for in other areas like instruction, assessment, and instructional delivery (Ismail et al., 2022), and most students' perceptions are negative (Elliott & Soifer, 2022; Hu & Min, 2023; Saura et al., 2022). The utilisation of AI in research cannot be achieved where students, academic staff, and stakeholders in higher institutions of learning are not concerned with the creation and integration of AI at various levels of instruction (Langran et al., 2020; Qin et al., 2020). The utilisation of AI in research requires that students acquire the skills and knowledge that will make them aware of its diverse applications as well as develop a positive perception of the role it plays in research (Seufert et al., 2020; Häkkinen et al., 2017).

The research effort remains unclear to this point. The researchers are not very exact on the level of utilisation of AI in research in Nigeria. This leaves the quality of work done in doubt because artificial intelligence has proven to be a veritable tool for excellent research outcomes. Admittedly, most of the existing studies look at the utilisation of AI for instruction in the educational circle and, specifically, its impact on international students' success (Wang et al., 2023) and the effectiveness of teaching (Almelweth, 2022). Other scholars have looked at AI in education and schools (Ahmet & Aydemirb, 2020; Chen et al., 2023), predicting the impact of AI on performance (Khan et al., 2021), and the challenges of AI for teachers (Ismail et al., 2022). However, the utilisation of AI among students for academic research has not been extensively explored.

Recent studies have tried to bridge the gap. Scholars like Adiguzel, Kaya, & Cansu (2023) studied revolutionising education with AI: Exploring the transformative potential of ChatGPT. The Chubb, Cowling, and Reed (2022) study appears to be the most recent and closely related to this study. The study focused primarily on the effect of AI on research practices and culture, using areas such as thematic analysis and deductive analysis to uncover issues affecting university staff. Besides, the utilisation of AI in research involves the engagement of different AI tools such as SciSpace, Schoarlcly, Jenni AI, ChatPDF, Paperpal, Casper, Grammarly, QuillBot, Turnitin, Elicit, Lateral, ClioVis, Glasp, Audiopen, Search Smart, Consensus, and Mendeley, among others (Huang et al., 2023b; Nazaretsky et al., 2022; Adiguzel et al., 2023). Thus, students' use of artificial intelligence tools for instruction and assessment is not the same as that used for review of literature, summarisation of studies, data analysis, plagiarism checks, and report writing. Further still, Bingimlas (2009) noted that students' utilisation is basically based on their level of awareness, perception, and access to machines that can be useful in their research endeavours. Similarly, students' level of awareness cannot be unconnected to the single fact that AI tools were not emphasised in African universities as compared to other universities in the world. Most of these tools were used in most international universities (Mogavi et al., 2023), and the paucity of conversation regarding the application of AI is germane for a study of this nature (Agyemang et al., 2023). Therefore, this study is imperative to guide policymaking in Nigeria and other developing nations.

## Literature review

Previous studies have been conducted on the utilisation of artificial intelligence in higher education (Liang et al., 2021; Hwang & Tu, 2021). In fact, since the inception of ChatGPT in 2004, there has appeared to be a paradigm shift in the number of studies in relation to ChatGPT, generative AI, and higher education. Studies on AI in relation to higher education are many (Limna et al., 2023; Xames & Shefa, 2023; Crawford et al., 2023; Ifelebuegu et al., 2023; Popenici et al., 2023; Adarkwah et al., 2023). For example, the Rasul et al., (2023) study found that AI is beneficial to higher education students in that it helps facilitate adaptive learning, personalised feedback, assessment, support research and data collection and analysis, as well as

automated management services. The study also found that research in areas of higher education is affected by issues of reliability, limitations in skill acquisition, academic integrity, and falsification of information. What has not been explored extensively is the utilisation of AI tools by university students in research. Studies that exist are few. Gasaymeh (2018)'s study showed that students own laptops and smartphones, which provide easy access to the utilisation of ICT in educational activities. The study also found that students' utilisation of ICT is high. However, the use of ICT may not necessarily mean AI because there are so many ICT facilities that students utilise, which may not be artificial intelligence tools.

The awareness of students of AI tools in research has been an issue of great concern among scholars since a variety of tools are available in research. This is because how students perceive and utilise the research tools that AI provides is crucial in developing good and quality work that meets global standards. Similarly, the awareness of AI tools among students is crucial to measuring how well they are prepared to utilise emerging technologies that are impactful in society. Gradually, students' perception of AI is changing. In fact, most students who were hostile to AI tools have gradually understood the importance of AI applications in their research studies and expressed optimism about AI assistance in various disciplines (Li, 2020; Miranty & Widiati, 2021; Fahmi & Cahyono, 2021). Kelly et al. (2023)'s study found that awareness differs across subgroups and disciplines. Other studies have also shown that students' level of awareness is very high, especially with ICT and manipulation of the social media space (Dessy Harisanty et al., 2022; Khanagar et al., 2021). Yelena et al. (2022)'s study found that students' level of awareness is low. The mixed-methods research demands that we provide empirical evidence that will further assist in decision-making. It is imperative that a further study that will provide more explanation for these lessons be provided, especially in Africa, where AI is still not adequately utilised among students.

Students' perceptions of AI tools in research have generated diverse opinions, primarily because many students express ethical concerns regarding the integration of AI within educational settings (Kung et al., 2023). However, research indicates a positive reception among students regarding the use of AI in research writing. This positive outlook is closely linked to their acknowledgment of the user-friendly interface of these tools and their capability to furnish additional materials that facilitate a deeper understanding of the subject matter studied (Arguson, et al., 2023). Other researchers have also focused on the awareness of the efficiency of AI in educational setups (Liang et al., 2021; Hwang & Tu, 2021; Ouyang et al., 2022; Chu et al., 2022). These studies found that AI has been very applicable in online higher education in terms of predictive performance, improvement of learning experiences, and automated assessment. Others still noted that AI in higher education is basically in areas like assistive technology, predictive modelling, content analysis, and image analytics (Yang, 2022; Hinojo-Lucena et al., 2019). The application of AI in research is minimal, and there is no universal agreement among scholars on the nexus between awareness and perception of the utilisation of AI tools in academic research. Almaraz-López, et al. (2023) studies

found that students are aware of the impact of AI and are willing to utilise it in education. However, the study was silent on their perception of AI tools. Syed and Al-Rawi (2023) found that 73% of university students have knowledge of AI, 69.4% thought it was applied only in health care, and 57.3% were aware of the impact of AI, but perception of AI was found to relate positively with year of study and nationality. Moreover, the study aligns with previous studies that attempted to show that when students perceive that the tools are suitable for their development, it enhances their utilisation at any level (Lund et al., 2023). These insights, therefore, underscore the connection between awareness, perception, and utilisation of AI tools.

However, gender studies have been carried out in different studies, especially as it concerns ICT usage and social media engagements (Owan et al., 2023). Most of the studies found gender to be non-significant in respect of awareness, perception and utilisation of technology. Contrary to this, there are other studies that do not state that the level of awareness of AI among students is significantly different between male and female students. Alimi et al.'s (2021) study found that the majority of tertiary institution students are not aware of the application of AI in learning and research and that both male and female students' levels of awareness of the use of AI are not different. Agyemang et al. (2023) found that 50 academics confirmed minimal awareness of ChatGPT. The findings of the study could be connected to the perception many students and staff have concerning AI, which has affected their utilisation. Syed and Al-Rawi's (2023) study found that student's level of awareness of AI is high and that they hold a positive perception about the concept, benefits, and implementation of AI tools in research. However, the negative challenge that some of the participants hold is basically a function of the manipulation of the tools, which they perceive as relevant but possess inadequate skills to operate.

The increasing utilisation of artificial intelligence has escalated the awareness of students, both male and female, but male students are often identified as being more aware than female students (McGregor et al., 2017; Odigwe & Owan, 2020). This concurs with previous studies conducted in Africa that posit that male students are more aware than female students of their engagements on the internet and utilisation of instructional technologies (Owan et al., 2023). Another study found that male students do not differ from female students in their perception of AI tools in research (Syed & Al-Rawi, 2023). In terms of utilisation, there are confusing reports of genders that use more ICT than others (McGregor et al., 2017). This is because there are some studies that tend to establish that in surfing the net, engaging in media charts, and being present in the cyberspace, no significant difference exists between males and females (Mesagan et al., 2022). The perception of students towards AI tools may concern data privacy and ethical implications. Further studies are necessary to examine these concerns so as to help students embrace emerging technologies.

In Nigeria, the study by Alimi et al (2021) revealed that students' level of awareness of AI tools is high. However, gender differences do not exist in awareness and utilisation of AI tools for learning. It is evident that most studies have

been able to establish the different rates of AI tool adoption in different academic research. However, there remains a need for a deeper analysis of the various factors that influence these discrepancies. Understanding why certain attributes like gender and programme type are important is necessary to tailor efforts and interventions in order to bridge the gaps. The use of AI in academic research and activities is a novel idea, especially in Africa. Most of the students do not have any fundamental knowledge of AI in their training as it is not part of the curriculum. In most cases, students are exposed to ICT programmes that only cover Microsoft Office applications and a little programming. Therefore, their knowledge of AI tools and their applicability could differ by programme type and discipline (Kasneci et al., 2023). There are insufficient studies that examine the differential variations between undergraduate and postgraduate students in relation to the level of awareness, perception, and utilisation of AI tools in research.

Currently, there is a need for an in-depth examination of research activities in the era of AI. This is because there are so many AI tools that are valuable in research activities. Students' non-use of AI tools in research could have a serious effect on the quality of the research outcome. For example, most students claim that they have not seen adequate literature on a particular area of interest. This could probably be due to their perception or limited knowledge of the various tools that are applicable to research. The recent study that was carried out in Saudi Arabia by Syed and Al-Rawi (2023) on perception, awareness, and opinion towards AI was more descriptive and only attempted to provide first-hand information on the characteristics of students with respect to how they conceive AI in their studies. These studies approach the issues from a bibliometric perspective. No study has been done to evaluate the awareness, perceptions, and utilisation of AI tools in research using a multi-group analysis technique. The rationale is that most universities in Africa still utilise traditional methods for conducting research. The level of digital materials that are necessary and required for full application of AI is not yet available, and lecturers too may not be aware of the diverse AI tools that can facilitate quality and efficient outcomes.

It is this literature that has provided the basis for formulating the hypotheses to unearth the intricate relationship existing between variables when certain factors like gender and programme type are involved.

- i. The extent of student's awareness of AI tools in academic research is not significantly high.
- ii. Students' perception of AI tools in academic research is negative.
- iii. There is a low level of student's utilisation of AI tools in academic research.
- iv. Students' level of awareness does not have a significant direct effect on their perception to AI tools engagement in academic research.
- v. Students' level of awareness does not have a significant direct effect on the utilisation of AI tools

in academic research.

- vi. Students' perception does not have a significant direct effect on the utilisation of AI tools in academic research.
- vii. The relationship between awareness and utilisation of AI tools in research is not mediated by students' perception of AI tools.
- viii. The direct effect of awareness on perception of AI tools in research is not significantly different between male and female students.
- ix. The direct effect of awareness on utilisation of AI tools in research is not significantly different between male and female students.
- x. The direct effect of perception on utilisation of AI tools in research is not significantly different between males and females.
- xi. The mediating effect of perception on awareness on the utilisation of AI tools in research is not significantly different between male and female students.
- xii. The direct effect of awareness on the perception of AI tools in research is not significantly different between undergraduate and postgraduate students.
- xiii. The direct effect of awareness on the utilisation of AI tools in research is not significantly different between undergraduate and postgraduate students.
- xiv. The direct effect of perception on the utilisation of AI tools in research is not significantly different between undergraduate and postgraduate students.
- xv. The mediating effect of perception in linking awareness to the utilisation of AI tools in academic research is different between undergraduate and postgraduate students.

## Conceptual framework

The conceptual framework presents the interlinkage of the variables with each other pictorially. This is presented in Figure 1.

## Methodology

The study is based on the positivist theory of research, which relies heavily on quantitative approaches. The study adopted a cross-sectional survey design. The cross-sectional design was applicable in the study since it only attempts to uncover associations by gathering data at a point in order to provide insight into the nature of the relationships. The study focused basically on the association between perception, awareness, and willingness for students' utilisation of AI tools for academic research. No attempt was made to manipulate the variables since it is a non-experimental study.

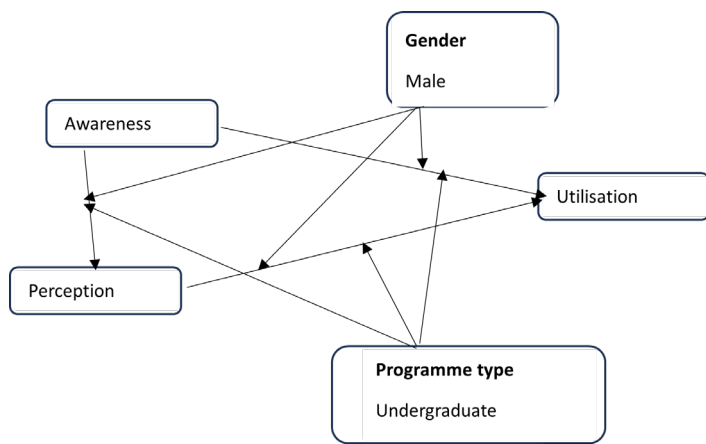


Figure 1: Conceptual model of the study.

The study participants were made up of final-year students in six universities in the study area, as well as master's and doctorate students who are in their second and third years of programmes respectively. These sets of students were selected because they are involved directly in writing their projects, theses, or dissertations. However, the eligibility criteria for selection were that the students' work must be quantitatively inclined since most of the tools in AI for researchers are both qualitative and quantitative. The study involved 5554 students. The demographic attributes of the students are that 3021 (54.39%) are male students, while 2533 (45.61%) are female students. In terms of programme type, 3232 (58.18%) are undergraduate students, while 2322 (41.82%) are postgraduate students. Similarly, 2102 (37.84%) are married, 2953 (53.17%) are single, and 499 (8.98%) are either divorced or widowed. The descriptive analysis also showed that 2611 (48.81%) are below 30 years, 2451 (44.13%) are between 30 and 50 years, while 492 (8.85%) are above 50 years.

### Instrument and measures

There are basically five measures in this study. These are gender, programme type, awareness, and perception of students, as well as the utilisation of AI tools in research. In this context, gender is defined as the biological characteristics that separate the male from the female. Programme type is the programme that the students are enrolled in, either as undergraduate students or postgraduate students. Awareness in this context is operationally defined as students' knowledge of the various AI tools that are used in research. Perception refers to the feeling or mindset that is either negative or positive that students hold about AI tools that are available to assist in research work. Utilisation of AI tools refers to the actual engagement of the plethora of available tools in doing research work by students. The instrument was divided into three sections. Section A was designed to provide demographic information about the respondents, such as gender, age, and programme type. Similarly, the section provided the opportunity for students to state their intention to participate in the study by ticking the check box provided in the section. This is after a cover letter was first made stating the objectives of the study, the confidentiality involved in the study, the need to provide objective responses, and how respondents' responses will

be protected from third parties. Options concerning gender, which was categorised as male and female, as well as type of programme, which was categorised as undergraduate and postgraduate, were provided in Section A.

Section B elicited responses on students' awareness and perceptions of the utilisation of AI in research works. This section, for clarity, was divided into two parts, such as awareness and perception. Awareness was measured with 7 items, with a sample item as "I am aware that AI tools can be used for literature reviews". Similarly, for perception of AI tools in research, 7 items were used for the study, and one sample item is "I sometimes feel that using AI for research is very unethical". These responses were to be obtained using a four-point Likert-modified option of strongly agree to strongly disagree. Section C was to elicit information on the utilisation of AI tools for research. This was done by listing the various AI tools that are necessary for research, and the respondents were to state the extent to which they have used the tools in their research work. On a linear scale of four-point response, the respondents were to indicate whether they had utilised the tool or not. Ten of the tools of AI that were featured include SciSpace, Schoarlcly, ChatGPT, Paperpal, Grammarly, QuillBot, Turnitin, Elicit, Consensus and Mendeley.

### Content validity

The content validity was carried out quantitatively using experts in diverse fields. Seven experts were selected from educational technology, measurement, and evaluation. These experts have high reputations in instrument development and analysis and have been in the field for the past ten years. Similarly, all those used in this study are professors in their respective disciplines. A total of 30 items were initially developed after an extensive review of the literature to identify what could constitute the domains of the variables. After initial screening, the initial pool of 36 items was reduced to 24 items that were considered suitable, relevant, and specific for the study. The decisions of the experts were based on an acceptable range of item-content validity indices (I-CVI) of 0.77 to 0.90 (for suitability), 0.78–0.99 (relevance), and 0.88–0.98 (precision). Items whose index was below 0.70 were reviewed for either clarity, relevance, precision, or both. This suggestion is in line with experts' opinions (Zamanzadeh et al., 2015). Similarly, for scale content validity indices (S-CVI), they ranged from 0.93–0.96, 0.91–0.94, and 0.90–0.98, respectively, for suitability, relevance, and precision. This helped to reduce the number of items from 24 to 20 based on the analysis as well as the comments by the experts in the comment form provided by the researchers to the evaluators.

### Preliminary analysis

The final draft was assembled that will be used for data collection. A total of 450 undergraduate and postgraduate students were selected for the pilot study. The selection of this number was based on the recommendation by various scholars that, in a survey, a ratio of 10:1 is enough to have a large sample (see Boateng et al., 2018). Thus, of the 24 items



for this exercise, 450 are considered adequate by this rule of thumb. The instrument was mailed to the respondents, and after five months (December 2022 – April 2023), the respondents had finished responding as expected, but only 5420 responses were obtained for the preliminary studies.

Exploratory factor analysis was carried out using the data obtained with varimax rotation based on maximum likelihood extraction techniques. A total of six factors were obtained from the initial analysis. However, some items were dysfunctional in that they loaded alone; some loaded in more than one factor, while others had factor loadings less than 0.30. These items were deleted, after which the remaining 15 items were loaded appropriately based on three factors. The three factors obtained, as presented in Table 1, explained a cumulative of 67.352% variance squared loadings. Each factor contributed to the total variance extracted. The utilisation of AI contributed 27.831% of the variance, the second factor (awareness of AI tools for research) contributed 22.964%, and the third factor (perception of AI tools for research) contributed 16.557% to the total variance. The KMO test of sampling adequacy yielded a coefficient of 0.794, while the Bartlett's test of sphericity yielded a significant result,  $\chi^2(105) = 3857.045$ ,  $p < .001$ , indicating that the correlation matrix was not an identity matrix and that the sample size of 420 was adequate or sufficient for the performance of factor analysis.

To establish discriminant and convergent validity, the study followed the suggestion of the Fornell-Larcker criterion (Fornell & Larcker, 1981) that relies mostly on the average variance extracted (AVE) and the composite reliability measures to determine these qualities. According to the scholars, where the AVE for each subscale is greater than 0.50, such measures are accepted as adequate for convergent validity, and where the square root of the AVE is greater than the inter-construct correlation coefficient of each of the subscales, it is established that discriminant validity exists. When these occur, it is always an indication that items could separate themselves from unrelated variables (Fresco et al., 2007; Patterson et al., 2005). The result in Table 1 presents the factor loadings of each item, the average variance extracted (AVE), composite reliability and discriminant validity of each factor.

Table 1: Exploratory factor analysis and internal structure of the scale to show dimensional evidence.

Items	M	SD	$\epsilon$	$\lambda$	$\lambda^2$	Construct attributes
UTI1	2.0690	.33491	.01634	.860	.740	AVE= .606 Discrim=0.778 $\alpha=.760$
UTI3	2.0905	.36741	.01793	.826	.682	
UTI5	2.0024	.55915	.02728	.769	.591	
UTI9	2.0548	.28377	.01385	.739	.546	
UTI8	1.5405	.64868	.03165	.736	.541	
UTI4	2.0929	.33627	.01641	.733	.537	
<b>SUM</b>	<b>11.8500</b>	<b>1.97934</b>	<b>.09658</b>	<b>3.927</b>	<b>3.637</b>	
AWR1	2.6976	.75145	.03667	.851	.724	
AWR4	2.7071	.64991	.03171	.824	.679	
AWR6	2.7262	.69350	.03384	.787	.619	
AWR3	2.7833	.73326	.03578	.703	.494	
AWR2	2.7833	.72012	.03514	.688	.473	
AWR5	2.7286	.76476	.03732	.627	.393	
<b>SUM</b>	<b>16.4262</b>	<b>3.24450</b>	<b>.15832</b>	<b>4.480</b>	<b>3.382</b>	
PER1	2.1071	.30966	.01511	.958	.918	AVE=.870 Discrim=.933 $\alpha=.779$
PER2	2.0905	.28720	.01401	.938	.879	
PER5	2.1310	.33775	.01648	.902	.813	
<b>SUM</b>	<b>6.3286</b>	<b>.88296</b>	<b>.04308</b>	<b>2.798</b>	<b>2.610</b>	

AVE=average variance extracted, Discrim=Discriminant validity,  $\alpha$ =Cronbach alpha

## Ethical consideration

The researchers had earlier explained in Section A of the questionnaire that participation in the study is voluntary. Options were also given for those who were not interested to tick appropriately. However, the study is a survey, and it does not cause any harm to subjects since none were subjected to any conditions. According to the Federal Ministry of Health (2007), ethical clearance can be waived. In spite of that, the respondents were made to tick the check box provided to indicate their willingness to participate in the study. In this way, consent was obtained by the respondents by writing to the researchers in the column provided in the questionnaire. The respondents were told that the responses would be used for publication in journal articles and that the information provided would be treated with a high level of confidentiality, to which no third party would have access without their consent. All those who had provided consent to this study were the respondents who finally responded to the questionnaires.

## Procedure for data collection

The data collection was done by sending a copy of the questionnaire electronically to the participants. This was done by engaging 40 research assistants who were financially motivated to support the team of researchers. The number of research assistants were high because of the large number of respondents in this study. The researchers were aimed at avoiding potential bias in the study. Each of the assistants was led by a principal author in this study. The researchers visited 6 universities in two geopolitical zones (South-South and South-East). The researchers were able to gain access to student union government leaders at the undergraduate and postgraduate levels. This helped to contact class representatives from various departments who are at their final year level and are writing research reports. These class representatives, based on their agreements, were added to the Telegram group created for this data collection. They were to share the instrument with their various class groups for the students to respond to it. They were, however, mandated to avoid sending the links to other forums that are not their class to avoid responses from those who do not constitute the frame of this study. The administration and collation of responses took about 8 months for the CVS file to be completed. However, a total of 5420 responses were downloaded, which indicated that this was the number that returned and took part in the study.

## Results/findings

Hypothesis One was tested using a population t-test to determine the level or extent of students' awareness of the use of AI tools in academic research. The result showed that the mean score of students' awareness of AI tool research is ( $M = 14.04$ ,  $S.D. = 2.81$ ) at a 95%CI [13.9678, 14.1175],  $t(5419) = 367.823$ ,  $p < .001$ . This showed that students' awareness of AI tools in research is significantly high. The alternate hypothesis is supported. Male students had a higher mean value ( $M = 14.60$ ,  $SD = 1.58$ ) of awareness of AI tools for research than the mean of female respondents

(M = 13.76, SD = 3.21), with a significant mean difference of -.840 and a 95%CI of [-.995, -.6880],  $t(5418) = 25.67$ ,  $p < .05$ . The study found that awareness of AI tools is stronger for male respondents than female respondents. Similarly, respondents who are postgraduate students have a stronger mean value (M = 15.42, SD = 2.87), compared to undergraduate students (M = 12.82, SD = 2.09), with a significant mean difference of -2.60 and a 95%CI of [-2.74, -2.47],  $t(5418) = -38.400$ . This showed that awareness of the utilisation of AI tools in research is stronger among postgraduates than among undergraduate students.

**Hypothesis Two: Students perception of the use of AI tools in research**

Hypothesis Two was tested using a population t-test to determine the extent of students' perceptions of the use of AI tools in academic research. The result showed that the mean score of students' perception of AI tool research is (M = 10.221, S.D. = 1.794) at a 95%CI [10.221, 10.317],  $t(5419) = 421.349$ ,  $p < .001$ . This showed that students' perceptions of AI tools in research are positive. The alternate hypothesis is supported. Male students had a higher mean value (M = 11.393, SD = 1.227) of perception of AI tools for research than the mean of female respondents (M = 9.712, SD = 1.769), with a significant mean difference of -1.678 and a 95%CI of [-1.773, -1.590],  $t(5418) = -36.188$ ,  $p < .05$ . The study found that male students have a more positive perception of AI tools than female students. Similarly, respondents who are postgraduate students have a relatively equal mean value (M = 10.22, SD = 1.750) compared to undergraduate students (M = 10.312, SD = 1.832), with a non-significant mean difference of 0.092 and a 95%CI of [-.092, -.188],  $t(5418) = 1.890$ ,  $p > .001$ . This showed that the perception of AI tools in research is similar among postgraduate and undergraduate students and is positive.

**Hypothesis Three: Students' utilisation of AI tools in academic research**

Hypothesis Three was tested using a population t-test to determine the extent of student's utilisation of AI tools in academic research. The result showed that the mean score of the extent of students' utilisation of AI tools research is (M = 20.932, S.D. = 5.67) at a 95%CI [20.797, 21.06],  $t(5419) = 303.776$ ,  $p < .001$ . This showed that students' utilisation of AI tools in research is significantly high. The alternate hypothesis is supported.

Female students had a higher mean value (M = 21.47, SD = 5.45) of utilisation of AI tools for research than the mean of male respondents (M = 19.845, SD = 1.769), with a significant mean difference of 1.626 and a 95%CI of [1.34, 1.91],  $t(5418) = 11.235$ ,  $p < .05$ . The study found that female students have higher utilisation of AI tools than male students. Similarly, respondents who are postgraduate students have a stronger mean value (M = 24.41, SD = 5.01) compared to undergraduate students (M = 17.82, SD = 2.39), with a significant mean difference of -6.59 and a 95%CI of [-6.79, -6.39],  $t(5418) = -62.76$ ,  $p < .001$ . This showed that the utilisation of AI tools in research is stronger among postgraduate students than among undergraduate students.

## Test of prediction

The test of prediction was carried out using partial least squares (PLS) structural equation modelling to determine the contribution of student perception and awareness to the utilisation of AI tools in research. Similarly, mediation analysis was carried out using students' perceptions of the relationship between awareness and utilisation of AI research tools. Figure 2 shows that student awareness and willingness which collectively explain 11.7% of the variation in student utilisation of AI tools in research  $R^2 = 0.117$ ,  $p < .05$ . Similarly, student awareness accounted for 3.8% of the variance in their student utilisation of AI tools in research,  $R^2 = 0.038$ ,  $p < .05$ . The f-square statistic shows that awareness and perception have significant effect sizes  $F^2 = .122$ , 95% [.08, .14],  $p < .001$ , 01. and  $F^2 = .095$ [.03,.07].07],  $p < .001$  in predicting student utilisation of AI tools in research.

The result for Hypothesis Four as presented in Table 3 indicates a significant direct effect of awareness ( $\beta = .19$ , 95%CI [.13,.23],  $t = 8.022$ ,  $p < .05$ ) on perception of AI research tools. Therefore, Hypothesis 4 was supported. The result for Hypothesis 5 as presented in Table 3 showed a significant negative direct effect of awareness ( $\beta = -.0167$ , 95%CI [-.22, -.06],  $t = 4.263$ ,  $p < .05$ ) on the utilisation of AI research tools in research. Thus, Hypothesis 5 was rejected by evidence. Similarly, on the direct effect of perception on the utilisation of AI tools in research (Hypothesis 6), the result is presented in Table 3.

Table 2: Population and independent t-test analysis of students level of awareness, perception and utilisation of AI tools in research by gender and programme types.

Variables	Extent of variables perception, awareness and utilisation				Gender				Programme type					
	N	M	SD	t-cal	Males		Females		t-cal	Undergraduate		Postgraduates		t-cal
					M	S.D	M	S.D		M	S.D	M	S.D	
Awareness	5420	14.04	2.81	367.82*	14.60	1.58	13.76	3.21	-10.42*	12.81	2.09	15.41	2.87	-38.40*
Perception	5420	10.26	1.79	421.34*	11.39	1.22	9.71	1.76	-36.18*	10.31	1.83	10.22	1.75	1.89
Utilisation	5420	20.93	5.07	303.77*	19.84	3.98	21.47	5.45	11.23*	17.82	2.39	24.41	5.01	-62.76*

*M=mean, SD=standard deviation, \*=Significant at .05 level*

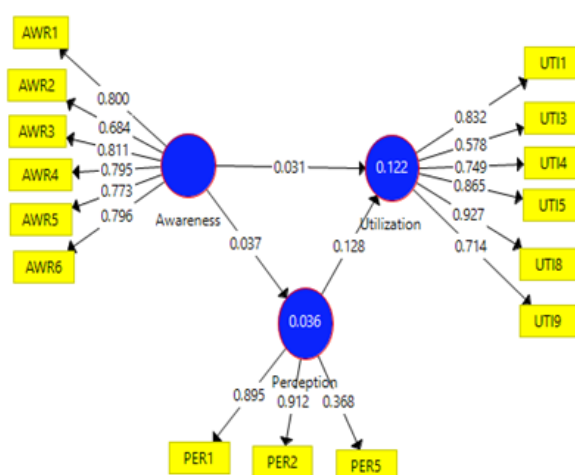


Figure 2: Structural Equation model connecting awareness, perception, and utilisation of AI in research.

Table 3 below showed that  $\beta = .341$ , 95%CI [.31, .37],  $t = 21.392$  ( $p < .05$ ). This shows statistical evidence that Hypothesis 6, which is on direct effect of perception on utilisation of AI tools in research, is rejected. Thus, the alternate hypotheses are supported for the three direct effects of awareness on perception, awareness on utilisation, and perception of utilisation of AI tools in research. The result of Hypothesis 7, as presented in Table 3, attempted to provide empirical evidence of the nexus between awareness and utilisation as mediated by perception and showed that ( $\beta = .065$ , 95%CI [.045, .081],  $t = 7.227$ ,  $p < .05$ ), which is an indication that perception provides a partial mediation between awareness and utilisation of AI tools in research. Thus, the null hypothesis is rejected.

Table 3: Direct and indirect effects.

Variables	$\beta$	95%CI	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>	Remarks
Awareness -> Perception	0.19	[.13, .23]	0.19	0.024	8.022	<.001	Rejected
Awareness -> Utilisation	-0.167	[-.22, -.06]	-0.162	0.039	4.263	<.001	Rejected
Perception -> Utilisation	0.341	[.31, .37]	0.342	0.016	21.392	<.001	Rejected
Awareness -> Perception -> Utilisation	0.065	[.046, .081]	0.065	0.009	7.227	<.001	Rejected

*M=mean, SD=standard deviation, CI=Confidence interval*

### Gender differences in the nexus between the independent and dependent variables

A multi-group analysis was carried out to determine the difference among respondents by gender on the nexus between awareness and utilisation, perception and utilisation, as well as awareness and perception of AI research tools among students. The result, as presented in Table 4, for Hypothesis 8 revealed that students' awareness significantly predicts their perception of AI tools for research positively for both males ( $\beta = .54$ ,  $t = 5.105$ ,  $p < .001$ ) and females ( $\beta = .11$ ,  $t = 2.78$ ,  $p < .001$ ), with the effect being stronger on males. The permutation test found a significant gender difference ( $\delta = -0.441$ ,  $p < .001$ ) in the prediction of awareness on students' perceptions of AI tools in research. Hypothesis 8, based on the result, was rejected. The results in Table 4 for Hypothesis 9 also showed that students' awareness significantly predicts their utilisation of AI tools in research positively for males ( $\beta = .11$ ,  $t = 0.929$ ,  $p < .05$ ) but negatively for females ( $\beta = -.20$ ,  $t = 9.772$ ,  $p < .001$ ), with the effect being stronger on the female students than the male students. The permutation test found a significant difference ( $\delta = .108$ ,  $p < .001$ ) in how awareness contributes to students' utilisation of AI tools for research more in females than males. Therefore, our hypothesis was rejected. Similarly, the result in Table 4 for hypothesis 10 showed that perception significantly predicted their utilisation positively for both males ( $\beta = .30$ ,  $t = 2.096$ ,  $p < .001$ ) and females ( $\beta = .31$ ,  $t = 18.409$ ,  $p < .001$ ), with the effect being relatively stronger in females than the male students. The permutation test found a non-significant difference ( $\delta = -.09$ ,  $p > .05$ ) in how perception contributes to students' utilisation of AI tools for research between males and females. Therefore, our hypothesis was sustained.

Hypothesis 11: Table 4 further shows that perception significantly mediated the nexus between students' awareness and utilisation of AI research tools, both positively for males ( $\beta = .17$ ,  $t = 1.980$ ,  $p < .001$ ) and females ( $\beta = .03$ ,  $t = 2.66$ ,  $p < .001$ ). The mediation effect was stronger for males than for female students. The permutation test reveals a significant difference ( $\delta = -.187$ ,  $p < .001$ ) in the mediation

Table 4: Multi group analysis based on gender.

Variables	B	95%CI	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
Males						
Awareness -> Perception	0.539	[.48, .59]	0.53	0.106	5.105	<.001
Awareness -> Utilisation	0.106	[-.21, .25]	0.08	0.114	0.929	>.05
Perception -> Utilisation	0.308	[-.27, .38]	0.274	0.147	2.096	<.001
Awareness -> Perception -> Utilisation	0.166	[-.16, .20]	.144	0.840	1.980	<.004
Females						
Awareness -> Perception	0.11	[.01, .17]	0.107	0.04	2.78	<.001
Awareness -> Utilisation	-0.202	[-.23, -.14]	-0.201	0.021	9.772	<.001
Perception -> Utilisation	0.311	[.28, .34]	0.312	0.017	18.409	<.001
Awareness -> Perception -> Utilisation	0.0340	[.02, .05]	0.034	0.013	2.660	<.008

Permutation test of differences

Paths coefficient	Paths coefficient					
	Baseline	Female	Males	$\Delta$	95%CI	<i>p</i>
Awareness -> Perception	0.19	0.106	0.547	-0.441	[-.01, .16]	<.001
Awareness -> Utilisation	-0.167	-0.22	-0.329	0.108	[-.08, .08]	<.001
Perception -> Utilisation	0.341	0.302	0.40	-0.098	[.09, .02]	>.05
Awareness -> Perception -> Utilisation	0.065	0.032	0.219	-0.187	[-.03, .05]	<.001

*M=mean, SD=standard deviation, CI=Confidence interval*

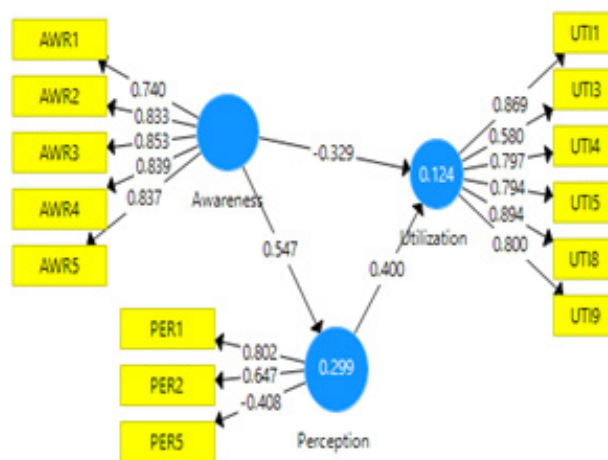


Figure 3a: Males.

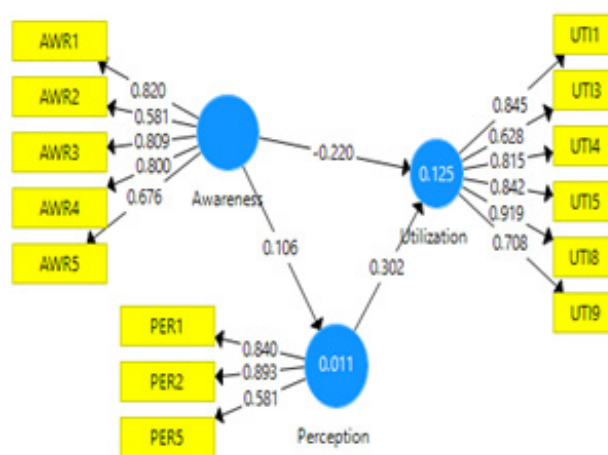


Figure 3b: Females.

effect of perception for both male and female respondents. Therefore, Hypothesis 11 was rejected. The result in Figure 3 further showed that awareness and perception, when combined, explain 12.4% of the variance ( $R^2 = .124$ ) in male students' utilisation of AI tools for research, while in

females, both variables, when combined, explain 12.5% of their utilisation of AI tools in research. Similarly, awareness explains 29.9% of the variance in male students' perceptions of AI tools, while for female students, it contributes only 1.1% of the variance in their perceptions of AI tools. This showed that students' awareness AI tools for research is stronger among the male students but lower among the female students in their perception of AI tools.

### Type of programme in the nexus between the independent variables and dependent variables

The result as presented in Table 5 for Hypothesis 12 revealed that students' awareness significantly predicts their perception of AI tools for research negatively for undergraduates ( $\beta = -0.27$ ,  $t = 2.444$ ,  $p < .001$ ), but positively for postgraduate students ( $\beta = 0.278$ ,  $t = 11.231$ ,  $p < .001$ ). The permutation test found a significant programme difference ( $\delta = -0.536$ ,  $p < .001$ ) in the prediction of awareness of students' perceptions of AI tools in research. Hypothesis 12, based on the result, was accepted. The results in Table 5 for Hypothesis 13 also showed that students' awareness significantly predicts their utilisation of AI tools in research negatively for undergraduates ( $\beta = -0.338$ ,  $t = 17.74$ ,  $p < .001$ ) but positively for postgraduates ( $\beta = 0.168$ ,  $t = 4.199$ ,  $p < .001$ ), with the effect being stronger among the postgraduate students than the undergraduate students. The permutation test found a significant difference ( $\delta = -0.506$ ,  $p < .001$ ) in how awareness contributes to students' utilisation of AI tools for research more in postgraduates than undergraduate students. Therefore, our alternate hypothesis was supported. Similarly, the result in Table 5 for Hypothesis 14 showed that perception significantly predicted their utilisation positively for both undergraduates ( $\beta = 0.146$ ,  $t = 2.85$ ,  $p < .001$ ) and postgraduates ( $\beta = 0.385$ ,  $t = 16.987$ ,  $p < .001$ ), with the effect being stronger in postgraduates than the undergraduates' students. The permutation test found a significant difference ( $\delta = -0.239$ ,  $p < .001$ ) in how perception contributes to students' utilisation of AI tools for research between undergraduates and postgraduates. Therefore, our hypothesis was supported.

Hypothesis 15: Table 5 further shows that perception significantly mediated the nexus between students' awareness and utilisation of AI research tools negatively for undergraduates ( $\beta = -0.04$ ,  $t = 4.067$ ,  $p < .001$ ) and positively for postgraduates ( $\beta = 0.103$ ,  $t = 9.497$ ,  $p < .001$ ). The mediation effect was stronger for postgraduates than for undergraduate students. The permutation test reveals a significant difference ( $\delta = -0.142$ ,  $p < .001$ ) in the mediation effect of perception for both undergraduate and postgraduate respondents. Therefore, Hypothesis 15 was rejected. The result in Figure 4a further showed that awareness and perception, when combined, explain 16.2% of the variance ( $R^2 = 0.162$ ) in undergraduate students' utilisation of AI tools for research, while in postgraduate students, both variables, when combined, explain 21.2% ( $R^2 = 0.212$ ) in their utilisation of AI tools in research. Similarly, awareness explains 7.2% ( $R^2 = 0.072$ ) variance among undergraduate students' perceptions of AI tools, while for postgraduate students, it contributes only 7.2% ( $R^2 = 0.072$ ) of the variance in their perceptions of AI tool research. This showed that students'

awareness and perception of the utilisation of AI tools for research are stronger among postgraduate students, but there is a higher relative equality in the contribution of awareness to perception between postgraduate and undergraduate students.

Table 5: Multigroup analysis based on type of programme.

Variables	$\beta$	95%CI	M	SD	t	p
Undergraduate						
Awareness -> Perception	-0.268	[-.34, -.14]	-0.248	0.11	2.444	<.05
Awareness -> Utilisation	-0.338	[-.37, -.30]	-0.343	0.019	17.74	<.001
Perception -> Utilisation	0.146	[-.01, .21]	0.138	0.051	2.85	<.001
Awareness -> Perception -> Utilisation	-0.040	[-.16, .20]	-.04	.010	4.067	<.001
Postgraduates						
Awareness -> Perception	0.268	[.21, .31]	0.268	0.024	11.231	<.001
Awareness -> Utilisation	0.168	[.08, .24]	0.167	0.04	4.199	<.001
Perception -> Utilisation	0.385	[.34, .43]	0.386	0.023	16.987	<.001
Awareness -> Perception -> Utilisation	0.103	[.02, .05]	0.104	0.011	9.497	<.000
Permutation test of differences						
Paths coefficient						
Paths coefficients	Baseline	UnderG	PostG	$\Delta$	95% CI	p
Awareness -> Perception	0.19	-.268	.268	-.536	[-.13, .13]	<.001
Awareness -> Utilisation	-0.167	-.338	.168	-.506	[-.07, .07]	<.001
Perception -> Utilisation	0.341	-.146	.385	-.239	[.06, .06]	<.001
Awareness -> Perception -> Utilisation	0.065	-.039	.103	-.142	[-.03, .04]	<.001

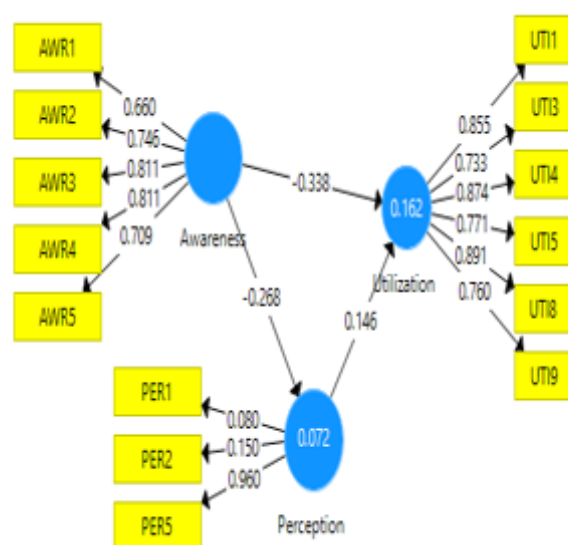


Figure 4a: Undergraduates.

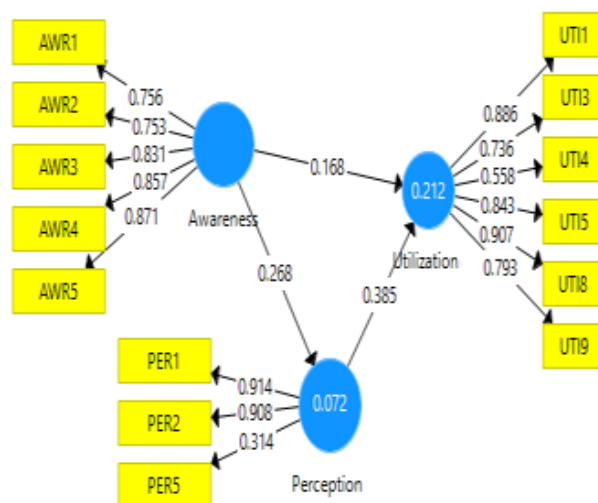


Figure 4b: Postgraduates.

## Assessment of outer model

Figure 2 is the baseline model that provides information about the outer loadings of the individual items to the latent factors. The items loaded appropriately to the various factors except for item UT3 (.578) in utilisation and PER5 (.368) for perception, which loaded lower than .70, which according to Memon and Rahman (2014) are desirable. However, the items were considered desirable since the other assessment criteria were well-fitted, as reported in the study (Götz et al., 2009). In Figure 3, the outer loading for gender was examined, and the results revealed that the loading of two items ranged from .408 to .894, while for female loading, it ranged from .581 to .919. For undergraduate students, item loading ranged from .080 to .960, and for postgraduate students, item loading ranged from .314 to .907. Importantly, some items loaded poorly into the latent construct, such as items PER 1 and PER 2 for undergraduates and PER5 for postgraduate students. However, these items were not deleted since removing them from the model affected the reliability of each subscale. These items for the undereducated may not be suitable for students in those programmes.

## Convergent validity

Convergent validity was established using the average variance extracted (AVE) for the measurement models in the study. It has been established that an AVE of 0.50 or above is adequate for achieving the convergent validity of a measure. The results in Table 6 revealed that all the variables, awareness, perception, and utilisation, obtained an AVE above .50, which is evidence that an AVE has been achieved. Similarly, for each group, gender, and type of programme as categorised, convergent validity was also achieved for males, females, undergraduates, and postgraduate students, but not for perception among undergraduate students.

Similarly, discriminant validity was assessed using the Fornell-Larcker criterion. It states that for discriminant validity to be achieved, the square root of the AVE for each construct must be greater than the coefficient of correlation among variables. The result in Table 7 presents empirical evidence of the discriminant validity of the constructs.

Table 6: Convergent validity of measures.

Variables	Gender			Type of programme	
	Baseline	Male	Female	Undergraduate	Postgraduate
Awareness	.847	.675	.552	.562	.664
Perception	.664	.509	.613	.417	.586
Utilisation	.899	.633	.638	.667	.634

Table 7: Discriminant validity of the measures.

Variables	Baseline		Males		Female		Undergraduate		Postgraduate						
Awareness (1)	0.77		.82		.74		.75		.81						
Perception (2)	0.21	0.76	.23	.71	.12	.78	.20	.51	.32	.76					
Utilisation (3)	-0.16	0.26	0.79	-.14	.33	.79	.29	-.13	.79	.18	-.16	.82	.28	.34	.79

*HTMT. HTMT values were less than 0.90 which is an indication that discriminant validity evidence for the population is achieved by gender and programme type.*

## Reliability

Two measures were used in determining the reliability of the measurement models, which are Cronbach alpha and composite reliability coefficient. Table 8 shows that all the reliability estimates are greater than .70. Therefore, the three subfactors—awareness, perception, and utilisation—had reliability coefficients across males, females, and undergraduate and postgraduate student subgroups.

Table 8: Composite and Cronbach alpha reliability estimates.

	Baseline		Gender				Programme type			
	α	CR	Males		Females		Undergraduate		Postgraduate	
			α	CR	α	CR	α	CR	α	CR
Awareness	.847	.885	.879	.912	.830	.859	.805	.864	.875	.908
Perception	.764	.780	.708	.890	.773	.822	.705	.708	.778	.780
Utilisation	.899	.908	.895	.910	.900	.912	.901	.923	.998	.916

*α = Cronbach alpha, CR = composite reliability*

## Discussion of findings

The result of the study showed that students in tertiary institutions are highly aware of AI in research. That is, they are aware of AI tools available for research. This level of awareness may relate to the rising popularity of AI among students in executing educational responsibilities. Studies have also shown that students spend a considerable amount of time with AI tools, especially ChatGPT, that have become very common in thesis writing, among others (Lattie et al., 2022). The study results also showed that male students have a stronger awareness of AI tools compared to female students. This aligns with previous studies showing that in Africa, the rising use of the internet for various purposes is more common among male students, probably a result of cultural segregation that exists. In Africa, it is common that most of the students who are women are always assigned home responsibilities, and this may reduce their time to have the required awareness of the plethora of tools that are useful for research purposes (Odigwe & Owan, 2020; McGregor et al., 2017). The findings of the study also showed that postgraduate students are more aware of AI tools used in research than undergraduate students. This finding is not unconnected to the fact that the postgraduate programme is research-oriented, and students are often exploring different avenues to get literature, knowledge of statistical tools, and how to beat plagiarism in their work compilation. This is in line with previous findings that have stated that postgraduate students are more exposed to ICT usage than undergraduate students (Ozimek & Bierhoff, 2016). This could also be due to the nature of the course work that they do, which to a very great extent, requires diverse tools that are optimally useful in their work organisation and the accumulation of information that provides explanations for areas of difficulty in their research expeditions.

The result of Hypothesis Two revealed that students' perceptions of AI tools in research are positive. The alternate hypothesis is supported. That is, irrespective of the wide general perception that many people have about AI in education, both students have a positive perception of AI in research. This is in line with previous studies (Liang et al., 2021; Hwang & Tu, 2021) that have evidenced that AI is useful in education for the purposes of programmed

learning, assessment, data collection, and self-tutoring (Ouyang et al., 2022). This relevance of AI in the educational circle and the assistance that it may have provided may be the reason why they hold a positive perception of tools in research. The result further showed that male students have a stronger perception of AI tools in research than female students. The findings may be due to the fact that, given their level of awareness of what AI tools are used for, they are more comfortable using AI to perform several tasks in research than the traditional sources of knowledge that most of the female students consider tedious and boring. This study contradicts previous findings that posit that female students are more social media-oriented than male students (Gil-Clavel & Zagheni, 2019; Oberst et al., 2016). The consistency in the dominance of male students in terms of awareness and perception of AI calls for further research to provide more explanation on the rationale for these differences, given that students of both gender need these tools equitably for their research engagement. More so, the result for the type of programme showed that both postgraduate and undergraduate students have a positive perception of AI tools in research. Both groups may have understood the relevance of these tools in their academic and research engagements, and thus, they need to develop a positive perception of them to utilise them adequately. The result aligns with a few previous studies that showed that students' perceptions of ICT at both the undergraduate and postgraduate level are positive, given that it is necessary for maximum results in research expeditions (Zhao, 2022; Kurniati & Fithriani, 2022).

The result showed that students in Nigerian tertiary institutions utilise AI tools for research with very high esteem. The extent of these findings could be because of the wide applicability of technology in research in universities and the educational system. This result is not surprising because it is common that in all the iPhones, laptops, and other gadgets that are held by students, these apps and tools are installed since students are widely tutored by different social media spaces, among other platforms. This helps them to utilise the various platforms and tools that have become very beneficial in editing, paraphrasing, and providing plagiarism checks. Similarly, there may be other factors within the environment that may instigate students' high level of utilisation of AI tools in research. These include the rising importance of AI in instruction and self-directed instruction, among others (Almaraz-López et al., 2023).

Similarly, the general rising awareness of students about the importance of AI tools in education and the reports about their easy accessibility for information found in several useful products may be another reason for students' high utilisation of these tools in research. However, contrary to expectations as defined by the cultural roles women play in Africa, women have a stronger level of utilisation compared to men. This contradicts earlier findings that male students are more involved in ICT usage compared to women (Christoph et al., 2015; Syed & Al-Rawi, 2023). The findings of the study may be connected to the fact that awareness is not utilised. One may be aware of AI tools but may lack the expertise and skills to utilise the various tools they are aware of. This could account for the differences because most of the female students in Nigerian universities hold sophisticated

phones and laptops that most of the male students do not have access to. These may have boasted more utilisation of the AI tools in research than the male students. The study has further implications, and other researchers could also carry out similar studies to provide explanations as AI is becoming more relevant in the education system and academic research. The study results further showed that postgraduate students utilise AI tools in research more than undergraduate students. This is also not surprising because those who are in their master's and doctorate studies are more involved in the search for knowledge, either to expand their knowledge base of the variables selected for the study or to increase the weight of the evidence from previous studies. This keeps them perpetually utilising tools associated with AI in their pursuit of research quality. The findings are similar to those of previous studies (Utami et al., 2023).

The result of this study revealed that students' awareness of AI tools significantly predicts their direct effect on their perception of AI tools in research. The outcome of the findings could be that the knowledge students have about AI determines what they perceive about AI. Most students' knowledge about AI is negative. First, most students have been made to believe that AI tool utilisation makes them redundant and unable to think for themselves; therefore, they become very lackadaisical in their engagement with AI tools in research. They perceive the use of AI to be negative and thus limit what they should have done with such development. Similarly, the direct effect of awareness on perception could also be connected to the fact that students who may have a low level of awareness may not perceive AI tools in a good light. The multi-group analysis further showed that male students have stronger awareness when compared to female students in their perception of AI tools. This is due to the cultural differences that have been established concerning male and female roles in Africa (Christopher et al., 2015; Syed & Al-Rawi, 2023).

The result of this study revealed that students' perceptions of AI tools significantly predict the direct effect on their utilisation of AI. Students who hold a positive view about the relevance of AI and its applicability in research will optimally utilise it in order to produce quality research work. More so, when students' perception is positive, they believe so much in those tools, given that the tools may help them access materials, paraphrase their works, and carry out editing. However, when students have a negative perception of AI tools, they see it from a moral perspective and may not use the tools as much. The outcome of the findings could be possible in that the perception one holds about an object determines the utilisation of such objects. The multi-group analysis further showed that male students are not different from female students in their perception of AI tool utilisation. This is due to the cultural differences that have been established about male and female roles in Africa. (McGregor et al., 2017; Odigwe & Owan, 2020).

The result of this study revealed that students' awareness significantly predicts the direct effect on their utilisation of AI tools in research. The rationale for the study could be that students cannot use what they don't have knowledge of. The more they are aware of the various AI tools that can

help their research work, the more they utilise them for that purpose. Secondly, students who, on a daily basis, come into contact with tools that aid them in carrying out one research work or another are more inclined towards its utilisation. This is because they have seen the relevance in the quality of work that the tools facilitate them to produce. Therefore, they may devote more time to engaging these facilities and tools for optimum research outcomes. The nexus between awareness and utilisation of AI tools was also found to be stronger from the male side than the female side. This result is not unconnected to the fact that previous studies have already stated that male students are stronger in ICT compared to female students (Odigwe & Owan, 2020; Owan et al., 2021) in surfing, downloading, and printing materials. Similarly, with respect to the type of programme differences in the relationship between awareness and utilisation, the findings further showed that postgraduate students and undergraduate students are both aware of the relationship that exists in their utilisation of AI tools. This is because of the rising level of technology in the educational sector and students' introduction to ICT at both the undergraduate and postgraduate levels. This introduction has raised their awareness of the use of AI tools for research by both students. The findings are also in line with previous studies that have stated that students' discipline and programmes are relevant in their utilisation of AI tools based on their awareness and perception (Wang, 2022).

The result of hypothesis that focuses on the variance explained in both male and female students based on their collective contribution of perception and awareness on the utilisation could be explained from diverse perspectives. First, the total variance was higher in male students than the female students. This could be due to the fact that societal expectations may influence the way male and female students engaged in technology and develop awareness and perception of AI tools. This perception may be tied to their traditional roles that may help male students be more aware of these tools and invariably affecting their perception. More so, the educational environment may contribute to the different levels of awareness and understanding of the use of AI tools. There are some curricula that expose more males to AI-related tools than the female students. In this sense, the cultural background where female students are restricted to certain activities may also play a role in these differences in awareness and perception that favours more males. This is similar to the outcome of the study that was conducted by Owan et al. (2023).

The result of the nexus between awareness and utilisation of AI tools as mediated by perception showed that perception has positively but significantly mediated the link between awareness and utilisation of AI research tools. This result may be because perception is very important, even if the student is aware of the relevance of AI in research. Many students are often concerned about what the outcome of using AI will hold for them and their future. This may also be a result of private concerns, as most students are made to believe that AI tools can make them redundant and less human if used for various purposes. This supports the findings of previous studies documenting that AI tools are suited for research (Miranty & Widiati, 2021; Fahmi & Cahyono, 2021). Similarly, the multi-group analysis further showed that male students,

who are stronger than female students in their awareness of AI tools, are also utilising them for their research. The findings are not far from showing that gender attitudes towards ICT are different. Most male students, especially now that the use of technology is diverse, are engaged deeply in it, especially in Africa, for diverse purposes. This may be the reason for the differences.

### **Limitations of the study**

The study has some limitations, especially when the interpretation of the results is involved. First, the study was carried out only in universities in Nigeria, and this may affect the generalisation of the study to other institutions like monotechnics, colleges of education, and colleges of health technology that were not incorporated into the study. Further studies are important for cities and institutions in the country. Secondly, the study was purely a survey that involved self-reports, which are not without personal biases and prejudices in the pattern of responses. Observation methods or interview methods could be more accurate. Finally, the use of a longitudinal method rather than a cross-sectional survey will have been more appropriate to comprehend students' awareness and perception of AI tools in research over time. This will help us understand the changes that happen in a student's awareness and perception as they progress in their research activities and academic ladder. However, this is not to say that the findings of the study are useless, as they have helped to provide more insight into the level of awareness and perception of Nigerian university students regarding the utilisation of AI tools in research. It also highlights the differential effects of gender and type of programme shaping awareness, perception, and utilisation of AI tools. Further studies can be carried out to address the challenges identified in this study and to provide more explanation for the outcome of the study.

### **Conclusion**

The conclusion drawn from this study is that students in Nigeria are highly aware of and have a positive perception of AI tools in research. The result also showed that the utilisation of AI is high among students. Students' level of awareness has a direct and significant effect on perception and utilisation of AI tools, with male students and postgraduates having stronger awareness of AI utilisation. Furthermore, students' perceptions of AI have a direct effect on their utilisation of AI, with a differentially stronger variation for male and postgraduate students than female and undereducated students. The findings are germane in that they have helped in providing more insight into students' use of AI based on variables like awareness and perception of these emerging technologies in research. The study contributes to AI promotion in existing literature in education. Therefore, the institutional technology base should be improved to enable students' access to free Wi-Fi so as to utilise AI tools for maximum research outcomes. There is a need for massive reorientation and sensitisation programmes that will provide students with the opportunity to handle media tools as well as learn how to utilise various research tools, especially among undergraduate students

whose awareness and perception are still very weak. The study also contributes to existing studies in that students who hold negative perceptions about AI tools will be more interested in engaging them through open access to internet-based platforms for quality research studies.

## Acknowledgements

The researchers acknowledged all validators of the instrument for their active participation in the process. The class representatives of various programmes that formed WhatsApp platforms for this purpose are also appreciated. Similarly, the students who provided consent for this study are also well appreciated.

## References

Adarkwah, M. A., & Huang, R. (2023). Technology addiction, abduction and adoption in higher education: Bird's eye view of the ICT4AD policy in Ghana 20 years on. *British Journal of Educational Technology*, 54(6), 1484-1504. <https://doi.org/10.1111/bjet.13352>

Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionising education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), Article 429. <https://doi.org/10.30935/cedtech/13152>

Aggarwal, C. C. (2018). *Neural networks and deep learning*. Springer. <https://doi.org/10.1007/978-3-319-94463-0>

Agyemang, M. A., Amponsah, S., Micheal, W., Ronghuai, D.D., Ahmed T. B., Ahmed, H., & Huanhuan, W. (2023). Awareness and acceptance of ChatGPT as a generative conversational AI for transforming education by Ghanaian academics: A two-phase study. *Journal of Applied Learning & Teaching*, 6(2), 78-93. <https://doi.org/10.37074/jalt.2023.6.2.26>

Ahmet, G. & Aydemirb, F. (2020). Artificial Intelligence in education and schools. *Research on Education and Media*, 12(1), 13-21. <https://doi.org/10.2478/rem-2020-0003>

Alimi, A. E., Buraimoh, O. F., Aladesusi, G. A. & Babalola, E. O. (2021). University students' awareness of, access to, and use of artificial intelligence for learning in Kwara State. *Indonesian Journal of Teaching in Science*, 1(2), 91-104. <https://ejournal.upi.edu/index.php/IJoTis>

Alloghani, M., Al-Jumeily, D., Mustafna, J., Hussain, A., & Aljaaf, A. J. (2020). A systematic review on supervised and unsupervised machine learning algorithms for data science. In *Supervised and unsupervised learning for data science* (pp. 3–21). Springer, Cham. [https://doi.org/10.1007/978-3-030-22475-2\\_1](https://doi.org/10.1007/978-3-030-22475-2_1)

Almaraz-López, C., Almaraz-Menéndez, F., López-Esteban, C. (2023). Comparative study of the attitudes and perceptions of university students in business administration and management and in education toward artificial intelligence. *Education Sciences*, 13(6), 609. <https://doi.org/10.3390/educsci13060609>.

Almelweth, H. (2022). The effectiveness of a proposed strategy for teaching Geography through artificial intelligence applications in developing secondary school students' higher-order thinking skills and achievement. *Pegem Journal of Education and Instruction*, 12(3), 169-176. <https://doi.org/10.47750/pegegog.12.03.18>

Arguson, A., Maborang, M., & Paculanan, R. (2023). The acceptability of generative ai tools of selected senior high school teachers in schools divisions office-manila, philippines. *Journal of Artificial Intelligence, Machine Learning and Neural Network*, 3(6), 1-10. <https://doi.org/10.55529/jaimlInn.36.1.10>

Bingimlas, K. A. (2009). Barriers to the successful integration of ICT in teaching and learning environments: A review of the literature. *Eurasia Journal of Mathematics, Science and Technology Education*, 5(3), 235-245.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioural research: A primer. *Frontiers in Public Health*, 6, 149. <https://doi.org/10.3389/fpubh.2018.00149>

Bonk, C. J., & Wiley, D. A. (2020). Preface: Reflections on the waves of emerging learning technologies. *Educational Technology Research and Development*, 68(4), 1595–1612. <https://doi.org/10.1007/s11423-020-09809-x>

Chen, Y., Jensen, S., Albert, L. J., Gupta, S., & Lee, T. (2023). Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Information Systems Frontiers*, 25(1), 161-182. <https://doi.org/10.1007/s10796-022-10291-4>

Chiu, T. K. F., Meng, H., Chai, C. S., King, I., Wong, S., & Yeung, Y. (2022). Creation and evaluation of a pre-tertiary Artificial Intelligence (AI) curriculum. *IEEE Transactions on Education*, 65(1), 30-39. <https://doi.org/10.1109/TE.2021.3085878>

Christopher, M. S., Goerling, R. J., Rogers, B. S., Hunsinger, M., Baron, G., Bergman, A. L., & Zava, D. T. (2016). A pilot study evaluating the effectiveness of a mindfulness-based intervention on cortisol awakening response and health outcomes among law enforcement officers. *Journal of Police and Criminal Psychology*, 31, 15-28.

Chu, H., Tu, Y., & Yang, K. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, 38(3), 22–42. <https://doi.org/10.14742/ajet.7526>

Chubb, J., Cowling, P., & Reed, D. (2022). Speeding up to keep up: Exploring the use of AI in the research process. *AI & Society*, 37(4), 1439-1457.

Clark, D. (2020). *Artificial intelligence for learning: How to use AI to support employee development*. Kogan Page Publishers.

Cohen, I. L., Liu, X., Hudson, M., Gillis, J., Cavalari, R. N., Romanczyk, R. G., ... & Gardner, J. M. (2017). Level 2



- Screening with the PDD behavior inventory: Subgroup profiles and implications for differential diagnosis. *Canadian Journal of School Psychology*, 32(3-4), 299-315. <https://doi.org/10.1177/0829573517721127>
- Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3), 2. <https://doi.org/10.53761/1.20.3.02>
- Dessy Harisanty, E., Variant, A., Tesa, E., & Aji, A. (2022). *Leaders, practitioners and scientists' awareness of artificial intelligence in libraries: A pilot study*. <https://www.emerald.com/insight/0737-8831>.
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Elliott, D., & Soifer, E. (2022). AI technologies, privacy, and security. *Frontiers in Artificial Intelligence*, 5, 826737. <https://doi.org/10.3389/frai.2022.826737>
- Fahmi, M. A., & Cahyono, B. Y. (2021). EFL students' perception on the use of Grammarly and teacher feedback. *JEEES Journal of English Educators Society*, 6(1), 18–25. <https://doi.org/10.21070/jees.v6i1.849>
- Federal Ministry of Health (2017). *National health research ethics committee*. <http://www.nhrec.net>
- Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). *Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers*. <https://doi.org/10.1101/2022.12.23.521610>
- Gasaymeh, A. (2018). A study of undergraduate students' use of information and communication technology (ICT) and the factors affecting their use: A developing country perspective. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(5), 1731-1746.
- Gil-Clavel, S., & Zagheni, E. (2019). Demographic differentials in Facebook usage around the world. *Proceedings of the International AAAI Conference on Web and Social Media*, 13(01), 647-650. <https://doi.org/10.1609/icwsm.v13i01.3263>
- Götz, D., Laurent, P., Lebrun, F., Daigne, F., & Bošnjak, Z. (2009) variable polarization measured in the prompt emission of GRB 041219a using IBIS on board integral. *The Astrophysical Journal*, 695(2). <https://arxiv.org/abs/0903.1712>
- Häkkinen, P., Järvelä, S., Mäkitalo-Siegl, K., Ahonen, A., Näykki, P., & Valtonen, T. (2017). Preparing teacher students for 21st century learning practices (PREP 21): A framework for enhancing collaborative problem solving and strategic learning skills. *Teachers and Teaching: Theory and Practice*, 23(1), 25–41. <https://doi.org/10.1080/13540602.2016.1203772>
- Hefernan, N. T., & Hefernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, 24(4), 470–497. <https://doi.org/10.1007/s40593-014-0024-x>
- Hinojo-Lucena, F. J., Aznar-Díaz, I., Cáceres-Reche, M. P., & Romero-Rodríguez, J. M. (2019). Artificial intelligence in higher education: A bibliometric study on its impact in the scientific literature. *Education Sciences*, 9(1), 51. <https://doi.org/10.3390/educsci9010051>
- Hu, Y., & Min, H. K. (2023). The dark side of artificial intelligence in service: The "watching-eye" effect and privacy concerns. *International Journal of Hospitality Management*, 110, 103437. <https://doi.org/10.1016/j.ijhm.2023.103437>
- Huang, A. Y., Lu, O. H., & Yang, S. J. (2023a). Effects of artificial intelligence-enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education*, 194, 104684. <https://doi.org/10.1016/j.compedu.2022.104684>
- Hwang, G. J., & Tu, Y. F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, 9(6), 584. <https://doi.org/10.3390/math9060584>
- Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIED) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching*, 6(2), 332-345. <https://doi.org/10.37074/jalt.2023.6.2.29>
- Ismail, C., Muhterem, D., Hanni, M. & Sanna, J. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66, 616–630. <https://doi.org/10.1007/s11528-022-00715-y>
- Kasneji, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneji, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
- Kelly, A., Sullivan, M., & Strampel, K. (2023). Generative artificial intelligence: University student awareness, experience, and confidence in use across disciplines. *Journal of University Teaching & Learning Practice*, 20(6), 12. <https://doi.org/10.53761/1.20.6.12>
- Kelly, M. S. (2023, January 26). ChatGPT passes exams from law and business schools. *CNN*. <https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html>
- Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021).

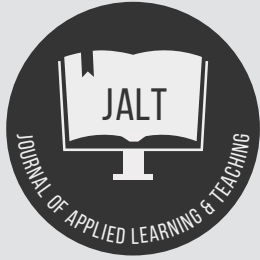
- An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments*, 8(1), 1-18. <https://doi.org/10.1186/s40561-021-00161-y>
- Khanagar, S., Alkathiri, M., Alhamlan, R., Alyami, K., Alhejazi, M., & Alghamdi, A. (2021). Knowledge, attitudes, and perceptions of dental students towards artificial intelligence in Riyadh, Saudi Arabia. *Medical Science*, 25(114), 1857–1867.
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLoS Digital Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
- Kurniati, E. Y., & Fithriani, R. (2022). Post-graduate students' perceptions of Quillbot utilization in English academic writing class. *Journal of English Language Teaching and Linguistics*, 7(3), 437–451. <https://doi.org/10.21462/jeltl.v7i3.852>
- Langran, E., Searson, M., Knezek, G., & Christensen, R. (2020). AI in Teacher Education. In D. Schmidt-Crawford (Ed.), *Society for information technology & teacher education international conference* (pp. 735–740). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/215821/>
- Lattie, E. G., Stiles-Shields, C. & Graham, A. K. (2022). An overview of and recommendations for more accessible digital mental health services. *Nature Reviews Psychology*, 1, 87–100. <https://doi.org/10.1038/s44159-021-00003-1>
- Li, K., Zhou, C., Luo, X. R., Benitez, J., & Liao, Q. (2022). Impact of information timeliness and richness on public engagement on social media during COVID-19 pandemic: An empirical investigation based on NLP and machine learning. *Decision Support Systems*, 162, 113752. <https://doi.org/10.1016/j.dss.2022.113752>
- Li, Z. (2022). Factors influencing students' continuous willingness to use e-learning platforms in higher education. *International Journal of Information and Communication Technology Education (IJICTE)*, 18(3), 1-11. <https://doi.org/10.4018/IJICTE.313424>
- Liang, J. C., Hwang, G. J., Chen, M. R. A., & Darmawansah, D. (2021). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2021.1958348>
- Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching*, 6(1), 64-74. <https://doi.org/10.37074/jalt.2023.6.1.32>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. <http://discovery.ucl.ac.uk/1475756/>
- Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. *Journal of the Association for Information Science and Technology*, 74(5), 570-581. <http://dx.doi.org/10.1002/asi.24750>
- McGregor, S. C., Lawrence, R. G., & Cardona, A. (2017). Personalisation, gender, and social media: Gubernatorial candidates' social media strategies. *Information, Communication & Society*, 20(2), 264-283. <https://doi.org/10.1080/1369118X.2016.1167228>
- Memon, A. H., Rahman, I. A., Abdullah, M. R., & Azis, A. A. (2014). Factors affecting construction cost performance in project management projects: Case of MARA large projects. *International Journal of Civil Engineering and Built Environment*, 1, 30-35.
- Mertala, P., Fagerlund, J., & Calderon, O. (2022). Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence (AI) and their implications for AI literacy education. *Computers and Education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caeai.2022.100095>
- Mesagan, F. O., Eseadi, C., & Omekwu, C. O. (2022). Influence of gender and expected competencies on access to and utilisation of cyberspace resources and services for research by postgraduate students. *Education and Information Technologies*, 27, 6157–6171. <https://doi.org/10.1007/s10639-021-10878-5>
- Miranty, D., & Widiati, U. (2021). An automated writing evaluation (AWE) in higher education. *Pegem Journal of Education and Instruction*, 11(4), 126–137. <https://doi.org/10.47750/pegegog.11.04.12>
- Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., ... & Hui, P. (2023). *Exploring user perspectives on chatgpt: Applications, perceptions, and implications for ai-integrated education*. arXiv preprint arXiv:2305.13114.
- Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, 53(4), 914-931. <https://doi.org/10.1111/bjet.13232>
- Oberst, U., Wegmann, E., Stodt, B., Brand, M., & Chamarro, A. (2016). Negative consequences from heavy social networking in adolescents: The mediating role of fear of missing out. *Journal of Adolescence*, 55, 51-60.
- Odigwe, F. N., & Owan, V. J. (2020). Academic staff personal variables and utilisation of ICT resources for research, teaching and records management in higher education. *The European Conference on Education 2020: Official Conference Proceedings*. <https://doi.org/10.22492/issn.2188-1162.2020.11>
- Okada, A., Whitelock, D., Holmes, W., & Edwards, C. (2019). E-authentication for online assessment: A mixed-method study. *British Journal of Educational Technology*, 50(2), 861–

- Ouyang, F., Zheng, L., & Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011–2020. *Education and Information Technologies, 27*, 7893–7925. <https://doi.org/10.1007/s10639-022-10925-9>
- Owan, V. J., Obla, M. E., Asuquo, M. E., Owan, M. V., Okenjom, G. P., Undie, S. U., Ogar, J. O., & Udeh, K. V. (2023). Students' awareness, willingness and utilisation of facebook for research data collection: Multigroup analysis with age and gender as control variables. *Journal of Pedagogical Research, 7*(4), 369-399. <https://doi.org/10.33902/JPR.202322235>.
- Owan, V. J., Abang, B. A., Idika, D. O., Etta, E. O. & Bassey, B. A. (2023) Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education, 7*(4), em2307. <https://doi.org/10.29333/ejmste/13428>
- Owan, V. J., Agurokpon, D., & Udida, U. J. (2022). *Curriculum restructuring and job creation among Nigerian graduates: The mediating role of emerging Internet applications*. SSRN 4009822.
- Ozimek, P., & Bierhoff, H.-W. (2016). Facebook use depending on age: The influence of social comparisons. *Computers in Human Behavior, 61*, 271–279. <https://doi.org/10.1016/j.chb.2016.03.034>
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*(1), 1-13. <https://doi.org/10.1186/s41039-017-0062-8>
- Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023). A critical perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning and Teaching, 6*(2), 311-331. <https://doi.org/10.37074/jalt.2023.6.2.5>
- Qadir, J. (2022). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education*. TechRxiv. <https://doi.org/10.36227/techrxiv.21789434.v1>
- Qin, F., Li, K., & Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology, 51*(5), 1693–1710. <https://doi.org/10.1111/bjet.12994>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., Santini, F. O., Ladeira, W.J., Sun, M.W., Day, I., Rather, A.R., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching, 6*(1), 41-56. <https://doi.org/10.37074/jalt.2023.6.1.29>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.29>
- Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2022). Assessing behavioral data science privacy issues in government artificial intelligence deployment. *Government Information Quarterly, 39*(4), 101679. <https://doi.org/10.1016/j.giq.2022.101679>
- Seufert, S., Guggemos, J., & Sailer, M. (2020). Technology-related knowledge, skills, and attitudes of pre-and in-service teachers: The current situation and emerging trends. *Computers in Human Behavior, 115*, 106552. <https://doi.org/10.1016/j.chb.2020.106552>
- Swiecki, Z., Ruis, A. R., Gautam, D., Rus, V., & Williamson Shafer, D. (2019). Understanding when students are active-in-thinking through modeling-in-context. *British Journal of Educational Technology, 50*(5), 2346–2364. <https://doi.org/10.1111/bjet.12869>
- Syed, W., & Al-Rawi, M. (2023). Assessment of awareness, perceptions, and opinions towards artificial intelligence among healthcare students in Riyadh, Saudi Arabia. *Medicina, 59*, 828. <https://doi.org/10.3390/medicina59050828>
- Vij, S., Tayal, D., & Jain, A. (2020). A machine learning approach for automated evaluation of short answers using text similarity based on WordNet graphs. *Wireless Personal Communications, 111*(2), 1271–1282. <https://doi.org/10.1007/s11277-019-06913-x>
- Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (ai) on international students in higher education: Generative ai, chatbots, analytics, and international student success. *Applied Sciences, 13*(11), 6716. <https://doi.org/10.3390/app13116716>
- Wang, Z. (2022). Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course. *Library Hi Tech, 40*(1), 80–97. <https://doi.org/10.1108/lht-05-2020-0113>
- Xames, M. D., & Shefa, J. (2023). ChatGPT for research and publication: Opportunities and challenges. *Journal of Applied Learning and Teaching, 6*(1), 390-395. <https://doi.org/10.37074/jalt.2023.6.1.20>
- Yang, Y., Liu, Y., Lv, X., Ai, J., & Li, Y. (2022). Anthropomorphism and customers' willingness to use artificial intelligence service agents. *Journal of Hospitality Marketing & Management, 31*(1), 1-23.
- Yelena, D., Rakhila, A., Botagoz, Y., Tulekova, M., Pigovayeva, N., & Samal, A. (2022). *Artificial intelligence awareness levels of students*. <https://doi.org/10.3991/ijet.v17i18.32195>.
- Yuan, S., He, T., Huang, H., Hou, R., & Wang, M. (2020). Automated Chinese essay scoring based on deep learning. *CMC-Computers Materials & Continua, 65*(1), 817–833. <https://doi.org/10.32604/cmc.2020.010471>
- Zamanzadeh, V., Ghahramanian, A., Rassouli, M., Abbaszadeh, A., Alavi-Majd, H., & Nikanfar, A.-R. (2015). Design and implementation content validity study: Development of an

instrument for measuring patient centred communication. *Journal of Caring Sciences*, 4(2), 165–178. <https://doi.org/10.15171/jcs.2015.017>

Zhao,X.(2022).Leveragingartificialintelligence(AI)technology for English writing: Introducing wordtune as a digital writing assistant for EFL writers. *RELC Journal*, 00336882221094089. <https://doi.org/10.1177/00336882221094089>

Copyright: © 2024. Ofem Usani Joseph, Iyam Mary Arikpo, Ovat Sylvia Victor, Nworgwugwu Esther Chidirim, Anake Paulina Mbu, Udeh Maryrose Ify and Otu Bernard Diwa. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Ethical concerns for using artificial intelligence chatbots in research and publication: Evidences from Saudi Arabia

Abu Elnasr E. Sobaih<sup>A</sup>

A

*Management Department, College of Business Administration, King Faisal University, Saudi Arabia, Hotel Management Department, Faculty of Tourism and Hotel Management, Helwan University, Egypt*

### Keywords

AI chatbots;  
artificial intelligence (AI);  
ChatGPT-4;  
ethical concerns;  
ethical dilemma;  
fake science;  
pseudoscience;  
scientific research.

### Abstract

Artificial intelligence (AI) conversational generative chatbots have drawn the attention of academics and have been increasingly used in the scientific research process since the inauguration of ChatGPT in November 2022. Despite growing research on AI chatbots' usage in research and publication, limited studies have deeply addressed the ethical concerns that arise from their usage. This research explores the perceptions of academics and their leaders regarding the use of AI chatbots in research and publication. It addresses the ethical dilemma and ethical approaches considered by academics and their leaders for shaping their decisions for the use or non-use of chatbots in scientific research. For these purposes, in-depth interviews were conducted with 21 academics/researchers and 11 leaders of scientific research in public universities in Saudi Arabia. The results of the thematic analysis confirmed that AI chatbots are extensively used in scientific research, albeit many researchers present their publications as their own work with no acknowledgement of the support from chatbots. The results showed ten interrelated ethical concerns, which would impact the growth of pseudoscience in developing countries if these concerns were not overcome. Hence, strategies for mitigating these ethical concerns are suggested. The research showed that academics often use chatbots based on a "utilitarian" approach, whereas most leaders consider the "virtue" or the "common good" approach for their concerns about chatbot adoption in scientific research. This research calls for policy and interventions from policymakers and other stakeholders about the responsible and ethical use of chatbots in research and publication.

### Correspondence

asobaih@kfu.edu.sa <sup>A</sup>

### Article Info

Received 9 February 2024  
Received in revised form 25 February 2024  
Accepted 29 February 2024  
Available online 7 March 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.21>

## Introduction

Chatbots go back to the 1960s when Joseph Weizenbaum developed the initial chatbot called ELIZA in 1966 (Weizenbaum, 1966). However, it was criticized because of limited knowledge and its inability to keep lengthy conversations and extract meaning from conversations. This was followed by updated chatbots such as PARRY in the 1970s, which addressed the shortcomings of ELIZA. It was more advanced and had a more controlling structure (Colby et al., 1972). More improvements were undertaken in chatbots after the integration of artificial intelligence in chatbots in the 1990s, introducing ALICE "Artificial Linguistics Internet Computer Entity", which relied on pattern matching, albeit it could not generate human-like responses (Heller et al., 2005). By 2000, chatbots were enhanced further to assist people in their daily activities and became available on Messengers such as Microsoft MSN. Thus, users were able to retrieve information from databases regarding news, weather and sports. This was recorded as a noteworthy improvement in human-computer interaction by gaining information from chatbots about daily activities (Adamopoulou & Moussiades, 2020). Voice assistance was then added to facilitate the use of chatbots and make voice commands, such as Google Assistant and Apple Siri (Adamopoulou & Moussiades, 2020).

Artificial intelligence (AI) conversational generative chatbots have witnessed a substantial transformation since November 2022 when ChatGPT was launched. ChatGPT has its place with the generative pre-trained transformer created by the American AI research lab OpenAI. OpenAI introduced ChatGPT as a fine-tuned, eco-friendly large language model (LLM), which can respond to numerous conversational prompts (OpenAI, 2022). Therefore, ChatGPT had one million users in the first five days and 100 million users in a few months. This made it the quickest-rising technology in history (Li, 2024). It continues to grow and reached over 180 million users in October 2023 (Li, 2024). ChatGPT has opened the door for several other recent AI chatbots, such as Google Bard (now Gemini), Bing Chat, and Ernie, which were developed to serve the needs of users for several purposes (Rudolph et al., 2023).

AI generative chatbots have drawn the interest of academics in higher education, and many of them started using this technology from day one (Al Lily et al., 2023; Al-Abdullatif, 2023; Lo, 2023). Rudolph et al. (2023) expected a "war of chatbots" for years to come, which is also anticipated to have a significant impact on higher education. Similarly, Hasanien and Sobaih (2023) predicted that the use of chatbots would have a massive impact on academic performance in higher education. Several scholars (see for example, Rudolph et al., 2023; Hasanien & Sobaih, 2023; Xames & Shefa, 2023; Ifelebuegu et al., 2023) have predicted a great change in higher education after the incorporation of AI chatbots. Chatbots were adopted for numerous purposes in higher education, e.g. academic writing, manuscript preparation, literature review, language editing, language translation, and statistical and data analysis (Hasanien & Sobaih, 2023; Gonsalves, 2023; Ifelebuegu et al., 2023; Joseph et al., 2024; van Dis et al., 2023; Xames & Shefa, 2023).

In the scientific research context, there was an increasing consideration from scholars about AI chatbot usage in scientific research (e.g. Hasanien & Sobaih, 2023; Gill et al., 2024; Manigandan & Sivakumar, 2024; Salvagno et al., 2023; Xames & Shefa, 2023). Previous research noted several benefits for the use of chatbots throughout the research life cycle for researchers (who undertake the research process and keenly disseminate it to other scholars or researchers), reviewers (who undertake the review process and judge the quality of submitted research) and editors (who make the decision about acceptance or rejection of publication based on certain criteria). However, some concerns were also noted in relation to authorship, referencing, academic integrity, plagiarism, information accuracy, copyright, privacy, transparency, and potential misuse of personal data (Chaka, 2023; Michel-Villarreal et al., 2023; Xames & Shefa, 2023). Despite these concerns, Xames and Shefa (2023) were optimistic about the use of scientific research in the near future. Hence, they argued that if these challenges were met, then AI generative chatbots would have a significant positive impact on scientific research.

Recent research on AI chatbot use in education, including scientific research (Rudolph et al., 2023; Hasanien & Sobaih, 2023; Xames & Shefa, 2023; Gill et al., 2024; Manigandan & Sivakumar, 2024) focused on exploring the perceptions and benefits of its adoption, drawbacks, or deficiencies in this technology. Ethical concerns for the use of AI chatbots were explored in general (Stahl & Eke, 2024). Recent studies (e.g. Popenici, 2023; Hasanien & Sobaih, 2023) have stressed that both scholars and higher education leaders must consider the ethical and intellectual implications of AI in education and research. However, to the best of my knowledge, there is no published research to date that investigated how the ethical decisions of users are shaped regarding the incorporation of AI chatbots in education and research. A previous study undertaken by Hasanien and Sobaih (2023) found that faculty and their students were more interested in the use of ChatGPT in education to support their study or work objectives, such as assignment preparation, language editing, translation, exam preparation and data analysis. In contrast, higher education leaders were more concerned about its use in education. Their concern was mainly because of its expected negative impact on learning outcomes, especially in the long term, such as critical thinking and other students' skills. This means that each key stakeholder has their own reasons or justification for their decisions. Notwithstanding this, it is important to understand how this decision meets ethical standards. What ethical approach was followed when making such a decision? How can it be enhanced to sufficiently consider the ethics and values of higher education in general and scientific research in particular? This research tries to answer these research questions. The research explores the perceptions of both academics and their leaders about AI chatbot usage in scientific research and their responses to the ethical dilemma of its use or non-use in scientific research. It explores ethical issues arising from the use of AI chatbots and ethical approaches considered by academics for shaping their decisions in the use or non-use of chatbots in scientific research. It is expected that a proper understanding of these concerns enables higher education to make the best use of technology in education and research, particularly AI chatbots.

## Literature review

### Chatbot use in scientific research

Since the launch of ChatGPT by the end of 2022, studies on the prospects and challenges for the use of AI chatbot in science and research are on the rise (Xames & Shefa, 2023; van Dis et al., 2023; Manigandan & Sivakumar, 2024; Kooli, 2023). Examples of these recent studies include the use of ChatGPT in research and publication (Xames & Shefa, 2023; Manigandan & Sivakumar, 2024; Salvagno et al., 2023). There was also a plethora of studies on the use of certain AI chatbots for various reasons, e.g. the use of "Quokka" chatbots for material science (Yang et al., 2024b); the use of "PLLaMa" chatbots for plant science (Yang et al., 2024a). Besides their use in education and research, AI chatbots were extensively used for other purposes, e.g. healthcare (Wang et al., 2023), cultural value proposition (Iodice et al., 2024), and banking and customer service (Le & Nguyen, 2015). There was also a plethora of studies (Rudolph et al., 2023; Waisberg et al., 2023; Ram & Pratima Verma, 2023; Cheong et al., 2023; Aiumtrakul et al., 2023) which focused on a comparison between various AI chatbots, i.e. ChatGPT, Bard, Bing, Baidu's Ernie, and their countless advantages and usages in different fields, particularly in education.

Regarding scientific writing and publication, most studies often focus on the opportunities of chatbots. For instance, Chen (2023) found that ChatGPT is beneficial for translation from Chinese to English. The same author found that ChatGPT is a great supporting tool in academic writing. However, Aydin and Karaarslan (2022) noted that the use of ChatGPT in writing could develop inadequate paraphrasing and plagiarism. Xames and Shefa (2023) reported several benefits of the use of ChatGPT in research and publication for researchers, editors and academics. For researchers, it could support them throughout the process of research writing and publication for the generation of ideas, the preparation of the manuscript, data generation and analysis, writing of the manuscript to full publication of the manuscript and language editing (Kim, 2023). It supports research in selecting the appropriate journal for publication and responding to incomprehensible comments made by reviewers. In addition, both editors and reviewers can make beneficial use of chatbots by making the right decisions about the submitted research and gaining more insights about it. AI chatbots can help both editors and reviewers make appropriate decisions about the manuscript and its revised version(s) (Xames & Shefa, 2023).

### The ethical dilemma of chatbot use in scientific research

Despite the well-documented opportunities for the use of chatbots in research and publication, a number of drawbacks or concerns were also noted (Hasanien & Sobaih, 2023; Xames & Shefa, 2023; Michel-Villarreal et al., 2023). These concerns are connected to the ethics of integrating AI into education and research. The first concern was associated with the authorship of chatbots (Lee, 2023; Lund & Naheem, 2023), like ChatGPT, in scientific publications since they have substantial contributions to the research. The question of "Can chatbots be considered a co-author in research?"

remains controversial to date (Xames & Shefa, 2023; Liebrezn et al., 2023). Some publishing bodies, such as Nature and Science, rejected the chatbots' co-authorship because they are not human beings and cannot take responsibility for their writing. In addition, copyright issues do not meet the current legal system (Kooli, 2023). Despite this, some articles were published with ChatGPT as a co-author (e.g. O'Connor & ChatGPT, 2022). Chatbots' co-authorship is also connected to copyright issues, accountability, and fairness. The question "Who holds the copyright for the text produced by chatbots?" remains unanswered and raises a concern about who the real co-author of the produced information is) Xames & Shefa, 2023(. In this case, who holds the legal responsibility is a key concern that should be addressed (Wang et al., 2023).

Another major ethical concern that emerged from the use of chatbots is the increasing risk of plagiarism and inaccuracies in research (Salvagno et al., 2023). Chatbots work on the data they were trained on, and hence, the responses taken from chatbots could be biased, inaccurate or have unintentionally plagiarized work because of algorithm bias. Chaka (2023) examined the accuracy of five AI content tools "GPTZero, OpenAI Text Classifier, Writer.com's AI Content Detector, Copyleaks AI Content Detector, and Giant Language model Test Room" to detect AI-generated text. The same author found that the five AI detectors were not efficient enough; hence, contribute to AI-generated plagiarism in publications. This opens the door for the growth of "pseudoscience" (Xames & Shefa, 2023). However, this sort of junk science could find room for publication in predatory publishers and their predatory journals, which do not rely on a solid peer-review process. Chaka (2024) recommended adopting contemporary AI detectors with traditional anti-plagiarism tools to ensure that the generated contents are human-written texts. It is not easy to differentiate between real and fake information without the eye of an expert in the field of science. Xames and Shefa (2023) revealed that ChatGPT, as the most common chatbots, has the tendency to suggest references that do not really exist. This raises a major concern about the validity and credibility of collected information from chatbots. It also raises another concern about the integrity of scientific research (Rudolph et al., 2023; Hasanien & Sobaih, 2023).

As discussed previously, chatbots rely on the scope and quality of the data they were trained on (Rudolph et al., 2023). Therefore, some chatbots, such as ChatGPT, were trained on data until 2021; hence, the gathered information may not be up to date or have a trainer bias (OpenAI, 2022). Although other chatbots, such as Google's Bard (now Gemini) and Bing Chat, are more updated, there are some concerns about inaccuracies of collected information from these sources that need human validation before full consideration of use in research (Rudolph et al., 2023). Another ethical concern is the fairness and equality in accessing information from chatbots as they are becoming paid services. However, many users from low-income nations will have limited access to these services (Salvagno et al., 2023). This imbalance in the service provision raises the digital gap between developed and developing countries, which is certainly not for the sake of research and science development (Xames & Shefa, 2023).

## Methodology

### The research approach

The study undertook a phenomenological approach to explore the decision of academics and their leaders about the use of chatbots in the scientific research process (Cilesiz, 2011). This kind of research approach helps the researcher/s delve deep into the experiences of academics, researchers, and leaders of scientific research. It helps in understanding how their decision about the use or non-use of AI chatbots in scientific research was formed (Neubauer et al., 2019). This study draws on the experience of academics in their use of chatbots and explores the ethical considerations relating to AI chatbot usage in their research and publication. This research approach also gave enough information about the perceptions of leaders of scientific research regarding the incorporation of chatbots in research and publication, especially how their decision is made and the ethical implications of their decisions.

### Data collection and sampling

The data were gathered from academics and leaders of scientific research in a sample of Saudi Arabian public universities. The data was collected from 21 academics and 11 leaders of scientific research who were deans and vice deans of scientific research in Saudi universities. Interviewees were accessed through personal networks at the different universities. This sample of research interviewees was adequate to reach data saturation (Fusch & Ness, 2015). There was slightly more participation from male academics (57 %) than females (43 %). Participants were nearly equally distributed to health sciences (23.8 %), engineering and computer sciences (23.8 %), agricultural and environmental sciences (23.8 %), and social sciences and humanities (28.6%). All the ethical issues were fully respected and adopted during the research process. This includes gaining approval from the authorized committee and interviewees' consent before the beginning of data collection, as well as hiding their name and identity (codes were used instead of name and job title) to protect their privacy.

### Data analysis

The interviews were voice-recorded or note-taken and transcribed after the interviews. This research used a thematic analysis for processing and analyzing the data (Braun & Clarke, 2012). Interesting points of data were coded, and themes were generated from the transcribed data. Three themes were developed. First, ethical concerns relating to chatbot usage in research and publication. Second, ethical decisions are made to use or not use chatbots in scientific research. Third, how ethical concerns relating to AI chatbot usage could be minimized and sorted.

## Results

### Theme 1: Ethical concerns relating to chatbot usage in research and publication

In-depth interviews with both academics and leaders of scientific research showed great value for chatbot use in research and publication. However, interviewees commented on ten key ethical concerns related to chatbot usage in research and publication. These ten concerns are privacy and confidentiality concerns, bias and inaccuracy of information, accountability and responsibility, authorship and licensing concerns, fairness and data accessibility, hallucinations concerns, recency of information, validation concerns, emergence of pseudoscience, and absence of human skills (see Table 1).

Table 1. The ten ethical concerns relating to chatbot usage in research and publication.

Concerns	Short description
1. Privacy and confidentiality	Collection of personal data by chatbots, which may be misused or unauthorized and accessed by third party, is a threat to users' privacy. Current privacy regulations do not fully fit with AI chatbots. The absence of a clear policy for personal data processing is a concern for confidentiality.
2. Bias and inaccuracy	Chatbots depend on the data they are trained with, and the techniques adopted by creators. Training data and techniques adopted by creators in chatbots could lead to biased and inaccurate information.
3. Accountability and responsibility	The quality of information relies on the developers or system creators who have to be accountable and act responsibly. Users also have to use this tool responsibly and recognize its drawbacks when using the information generated by chatbots.
4. Authorship and licensing	The authorship of outputs is a major concern. Who owns the authorship (e.g. system creators, users or chatbots themselves) or have the copyright and licenses of the outputs generated by chatbots.
5. Fairness and data accessibility	Researchers who are familiar with technology and those who could pay for some paid chatbots have more access to chatbots than other colleagues in many developing countries raising a concern about equality and fairness.
6. Hallucinations of chatbots	Chatbots produce hallucinations, including fake or inaccurate citations and unintentional plagiarism, which cannot be easily identified by many software.
7. Recency of information	Some chatbots provide data that is not up-to-date and may have algorithm bias. This raises concern about data accuracy and recency of information used.
8. Validation of information	Another key concern is the validity of generated information. The outputs of chatbots are not 100 % accurate and require expert assessment and validation before consideration in research or publication
9. Lack of human skills	Extensive use of technology, AI and chatbots has a negative effect on human interaction and interpersonal skills and may lead to self-isolation. Some interviewees were concerned about mental health because of the intensive use of chatbots.
10. Emergence of pseudoscience	The spread of inaccurate and biased chatbots' outputs without proper peer review in predatory journals and online rise pseudoscience sciences. Other scholars could build their work on this wrong information, which would raise a concern about the future of science.

The first concern raised by interviewees was related to the possibility of gathering personal data and chat history by chatbots. Interviewees were worried that their information may be misused, possibly for cyberattacking. The probability of using information of users in any purposes by chatbots or third party without their approval raises an important concern about chatbot usage in research and publication. This concern has also a legal aspect relating to data protection. It was argued that the current legislations and regulations related to data protection and privacy do not fit with AI and require an update to consider these aspects. Some comments about this issue are:

I am so worried about my personal data and searching history. They could be linked together and misused [Res12].

As I know, there is no national policy that addresses AI usage in education and re-search. Confidentiality of our researchers is a key concern for chatbot use in scientific research [Lead4].



The second concern was associated with bias and inaccuracy of output due to the bias and inaccuracy from data used to train the chatbots or the techniques adopted by system developers, which may lead to inaccurate and biased data. Interviewees have experienced wrong information generated by chatbots that are inaccurate and argued that they need validation before consideration in research or publication. Interviewees from all backgrounds, i.e. different disciplines, confirmed that they had found inaccurate data generated by chatbots, and hence, they raised a concern about data bias and inaccuracy. As one of the researchers from social science commented:

I have experienced several outputs that are not true. For instance, in my recent search, I found wrong assumptions about one of theories .... that has been updated but the output gave me the old assumption about this theory [Res14].

Another researcher [Res12] commented on the bias of the ChatGPT questioning, "Have you tried to debate with ChatGPT and assume that you are right, and the output is wrong?" He then answered himself and said, "I have done this a couple of times and ChatGPT changed the answer". This raises a concern about the accuracy of outputs that would be used to build scientific research and publication.

The third concern was related to the accountability and responsibility of both creators and users or academics, who must be responsible in their creation and usage of this supporting tool. Accountability of the system creators is a major ethical concern regarding the quality of output by chatbots. The quality of information used in scientific research cannot be questioned and has a significant impact on the integrity and outcomes. Responsible use is also a concern because some scholars use chatbot outputs with no acknowledgement of the source. They used AI chatbots to assist them in writing, developing their manuscript and generating responses to reviewers, but they provided this as their original work, which raises again concerns about accountability, responsibility and integrity. An academic commented that:

Recently, I received a comment from a reviewer about my manuscript submitted to one of the international journals, Q2 Scopus and ISI indexed, by the way. I asked ChatGPT about this, and I have used the output in my response to the reviewer and in my revision. The paper is now published. To be honest, I do not understand this point to date [Res7].

The above researcher who published this paper confirmed that this paper was published at the beginning of 2023 before most international journals used software to check AI use in publication. However, some international journals do not check AI to date and do not have the proper software for checking the use of AI in their published work. In addition, the contemporary AI detectors are not 100 % accurate and efficient in detecting AI-generated contents. This requires adopting AI contemporary detectors alongside traditional anti-plagiarism tools for making sure that the published text is generated by humans not AI tools.

The fourth concern is associated with authorship, copyright, and licenses. More specifically, leaders of scientific research were concerned about who owns the copyright of the chatbots' output. It could be the system creator, users of the system, the chatbots themselves, or the main source from which the trained data came. Again, many academics argued that they are using chatbots throughout their research process but do not really mention or acknowledge them in their work. This is because they believe it is an assistance tool and could be used but cannot be referenced or mentioned in the publication. An academic commented:

I use many AI chatbots in my research, such as ChatGPT and Bard, and I do not mind adding them as my coauthors or citing them in my article, but how can I do this? Policies of most publishers do not accept them. What shall I do? Do you think I should stop using them in my publication? [Res5].

In this case, the work was published as an original contribution by the authors, which again raises a major concern about research integrity. Regarding this point, one of the scientific research leaders commented:

Any leader in higher education or research institutions has to be concerned about the use of chatbots in scientific research and publication. We do not really think that any AI tool could be an author or hold the copyright. It could be an assisting or supporting tool. We have developed this tool to help us, not to be the author... Where is the academic and research integrity? [Lead1].

The fifth concern was related to equality in chatbot accessibility and usage. Interviewees argued that ChatGPT-4 is a paid source and hence many researchers from developing countries are not able to access this tool. Instead, they use the free version with limited features. This could be seen in the following comment by one of the international researchers who works in Saudi Arabia:

I can pay for any AI tool, but my colleagues in my home country could not really do this. They ask me to help them. I do this but they could not pay or have a communication with other scholars. Do you think this is fair or ethical? [Res11].

Another concern raised by interviewees is that some aged academics are not familiar with AI technology and found it difficult to use technology without assistance from other colleagues or a specialised unit at their institution, which does not exist in all institutions, where interviewees are currently working. The following comment by a young researcher explains this issue:

Not all researchers found it easy to use AI chatbots. Some of them do not really know about this because they are not familiar with the technology in general and AI in particular. Many of my professors are not aware of this, to be honest [Res18].

This means that despite AI's potential to bridge the technology gap, it is now contributing to widening the gap between digital natives and non-digital natives and between those from developing and developed countries, who can easily pay and access all chatbots.

The sixth concern was associated with hallucinations of chatbots and its tendency to imitate humans resulting in fake and inaccurate references and increasing rate of unintended plagiarized work. Interviewees commented about this arguing that work generated by ChatGPT was neither noted by "Ithenticate" nor many of AI detectors as a plagiarized work, though it is used by most of international journals. This requires the use of both tools in order to ensure the originality of the submitted work. They have submitted it to one of the international Scopus-indexed journals and published it recently. Another interviewee commented that although their university has a subscription to Turnitin, they have to pay additional fees to check AI use in their publication. Nonetheless, institutions have to subscribe to AI-checking services by Turnitin to allow their researchers access to such services without additional payment. This again raises a burden on some young and public universities in developing countries, where they have limited financial resources.

We work in a team, and some of us may use AI. Hence, we need to double-check before submission. We can pay for this service, but many other researchers in our home country could not do this. Again, do you think this is fair and ethical? [Res21].

The seventh concern was associated with the recency of the ChatGPT outputs, which is limited to 2021. Hence, concerns about the recency and accuracy of information remain under question. This requires the researcher to use other chatbots connected to the internet, such as Bing Chat and Bard (now Gemini), to double-check the updated information. This is connected with Concern Number Eight, relating to the validation of information given by chatbots for use in research and publication. There was a consensus among the researchers and leaders that the outputs of chatbots must be validated by researchers and experts in the discipline of research before publication. One of the leaders commented:

Validation of outputs is the responsibility of our researchers. We should not trust the outputs of AI chatbots without double-checking and peer reviewing [Lead3].

The ninth concern for the use of chatbots is the negative effect on human interaction and interpersonal communication skills with other academics and scholars. Some of the interviewees argued that since they started to use AI chatbots extensively, they had to limit their work with other scholars and work alone with AI tools' support. There were comments by several academics that their extensive use of AI chatbots may have a negative impact on their mental health. This could be noticed from the following comment:

I have extensively used technology, particularly AI chatbots, in my teaching and research. I prefer self-

isolation and less contact with other researchers, but I started to feel headaches and I am tired most of the time. I guess it could be due to my huge use of technology [Res19].

The last concern is critical, as many interviewees argued that the responsible use of chatbots with limited peer review for validation has led to the emergence and spread of pseudoscience. Some interviewees shared their experiences in using Bard and ChatGPT to develop their research articles and argued that most of their articles were developed with assistance from chatbots, and they published several research papers with the assistance of this tool. However, scientific research leaders were concerned about the outputs of this tool and confirmed that caution should be taken when disseminating this information, especially in international publications. Among the comments of scientific research leaders:

I am indeed worried that the irresponsible use of AI in research with no proper validation could spread junk science. Predatory journals would find room to publish such research to generate money from publication [Lead6].

We need to be careful about ChatGPT and other similar tools in publication. We need to take our role in managing this usage seriously as it could lead to the presence of pseudoscience that includes some assertions that are not really true or scientific [Lead2].

## **Theme 2: Strategies for mitigating ethical concerns relating to chatbots**

Interviewees agreed that the ethical concerns relating to chatbot usage in research and publication would be overcome if the proposed strategies for mitigating these concerns were adopted (see Table 2). There was consensus among concerned leaders and researchers that there is a quick need to develop a policy and guidelines on the responsible and ethical use of AI research and publication. The policy should maintain the privacy of users and protect their personal information. This policy should clearly explain the accountability and responsibility of each key stakeholder. Among the comments of both leaders and researchers:

There is no doubt that we are in a need for a policy organising this AI use in education and research. Our policymakers should supplement it with guidelines that address all associated concerns [Res9].

The increasing use of chatbots by researchers calls for rules that organise this usage in our institutions [Lead2].

The accountability and responsibility of creators ensures accurate and unbiased data when they develop the system. In addition, the responsible use of outputs by academics is important for a sustainable research process. It is the main responsibility of the academics/researchers to double-check the accuracy of the quality of outputs. Both leaders and researchers agreed that our academics or researchers must recognize the hallucination feature of chatbots and thus the outputs may not be accurate. Therefore, it is their

responsibility to validate the quality of outputs and double check the recency of the assumptions provided by chatbots. Additionally, the use of AI detectors alongside traditional anti-plagiarism programs are recommended to ensure the text is not AI-generated contents and human original work. Here are some examples of the interviewees' comments about this issue:

I think it is our responsibility as scholars to ensure that the information gathered from whatever source is accurate and scientific. We have to check the validity of collected information as we always do [Res9].

Our academics have to recognise the limitations of these AI tools and use them responsibly to advance their knowledge, not anything else [Res15].

Table 2. Strategies to deal with ethical concerns relating to chatbot usage in research and publication.

Concerns	Proposed strategies to mitigate concerns
1. Privacy and confidentiality	There should be a clear policy and guidelines for protection of users' information and maintaining their privacy and confidentiality with special consideration of AI tools. Personal data must not be used without the approval of the users.
2. Bias and inaccuracy	Academics and researchers should maintain awareness and undertake a critical review of chatbots' outputs to ensure their accuracy. Peer review remains the most valid approach to date for ensuring the quality and validity of the outputs.
3. Accountability and responsibility	Everyone from the key stakeholders (e.g. policy makers, creators and users or academics) should undertake their role responsibly and ethically. This should be part of the overall policy on IA incorporation in research.
4. Authorship and licensing	Policy makers and international research dissemination businesses have to establish a clear policy and guidelines on the AI outputs authorship, copyrights and licenses. The policy must address concerns relating to authorship, copyrights and licenses.
5. Fairness and data accessibility	Universities and research centers should facilitate responsible AI usage by bridging the digital divide and allowing up-to-date technology. A supporting AI unit could be established at each institution to assist academics in their work.
6. Hallucinations of chatbots	Recognizing the drawbacks of chatbots is important. Checking the validity of outputs is the responsibility of academics/ researchers. Acknowledgement of all sources, including chatbots, is important for ethical research.
7. Recency of information	Recognizing the limitations of the latest outputs by some chatbots is important for assigning the quality of outputs. Using various chatbots and databases is important to ensure the recency of information.
8. Validation of information	Peer validation and evaluation of the chatbots' outputs remain the key tools for assessing the quality of research and publication. Outputs should be reviewed by experts to ensure the quality of outputs. The use of AI detectors alongside traditional anti-plagiarism programs is recommended.
9. Absence of human skills	Quality of scientific research should balance the integration of AI in research and communication with other scholars for better and sustainable outcomes, especially in the long term.
10. Emergence of pseudoscience	Academic integrity and sustainability of scientific research are the responsibility of everyone from the key stakeholders (i.e. policymakers, university and scientific research leaders, AI system creators, and academics)

The policy should also ensure equality and fair accessibility to AI tools and chatbots that assist academics in their research. For example, the IT unit at each institution could provide support for academics about the effective use of AI chatbots in education and research. This includes providing training sessions, workshops and ongoing technical support. Many academics and their leaders, as can be seen from the following comments raised this point:

I think there should be an AI-supporting unit in this digital era of learning. This unit gives the required technical support for academics and students. Training programs and workshops on the effective use of AI in research could also be provided [Res4].

All Saudi institutions have an IT support unit, but I do not think that they provide support for AI use

in education or research. Why not integrate AI into their current responsibilities [Lead2]?

Interviewees suggested that policymakers of higher education and international research publishing businesses must publish clear guidelines regarding the authorship, licenses and copyrights of the chatbots' outputs used in research and publication. All these suggested activities would certainly make the best use of such technology. It should support research integrity and limit the dissemination of pseudoscience as approved by interviewees.

### Theme 3: Ethical decision to use or non-use chatbots in research and publication

Both academics and leaders of scientific research were asked about what shaped their ethical decision to use or not use chatbots in research and publication. Academics fully agreed about the use chatbots in their research. In their discussion about this point, they paid more attention to the benefits and opportunities that chatbots create for them as researchers, reviewers, and editors of research outputs. They focused on their comments on counting the advantages of chatbots, e.g. idea-generation for research, undertaking a literature review, translation, and proofreading service for non-native speakers, data analysis and final manuscript preparation for publication. They also commented on the chatbots in suggesting journals for submission and suggesting reviewers for the manuscript. Chatbots also supported them in their perpetration of responses to reviewers' comments. Academics argued that chatbots are a great supporting tool, which makes their research process easier. Chatbots also supported them in their review and editing of manuscript to make the right decision about the status of manuscript. This group of participants believe that chatbots provide more good than harm for their research and publication. This could be seen from the following comments:

The benefits of using AI chatbots are countless and unlimited. They are fantastic tools that made our research journey easier. We can use them in every step, and they give us quick and prompt responses. I think they are advancing scientific research forward [Res3].

Let's compare the advantages of ChatGPT with its disadvantages and find an answer to why we are using it. It made our research process simpler... I understand some people may talk about the inaccuracy of information, but we can validate this. Look at the numerous advantages, please [Res1].

Hence, this group of research participants focuses on the benefits or advantages to justify their use of AI in their research. They followed a utilitarian approach when making decisions about the use of chatbots in their research and publication, with little attention paid to the negative consequences.

The second group of participants, who are mainly leaders of scientific research, are concerned about the integration of chatbots in research and publication because their harm exceeds their benefits or good. This group of participants are concerned about academic integrity and the presence of fake and/or junk science because of chatbots' hallucinations as well as biased and inaccurate outputs, which could find room for publication in predatory journals or online without a proper peer review process. This group of participants were much concerned about the long-term impact of chatbots integration if they are used irresponsibly. Therefore, their decision was made based on virtue and common good approach that prioritise virtues such as integrity, equality, and accountability for the sustainability of research. This could be observed from the following comments:

We need to look at the long-term impact. Yes, AI chatbots give some support to our students and academics in learning and research. However, limited attention to the limitations of these tools and irresponsible use of these answer generators could produce fake science [Lead10].

I think caution should be taken when institutions decide to integrate such AI tools in learning or research. We have to look at both positive and negative impacts. I am not well-motivated to officially integrate them because I believe their negative impacts are undoubtedly higher than their positive impacts [Lead1].

## Discussion

Since the launch of generative AI chatbots, they have gained high acceptance from users in many fields including healthcare (Wang et al., 2023), customer service (Huang et al., 2024), and education (Rudolph et al., 2023; Al Lily et al., 2023; Al-Abdullatif, 2023; Hasanien & Sobaih, 2023; Xames & Shefa, 2023). In the same context, previous studies (see, for example, Xames & Shefa, 2023; van Dis et al., 2023) showed that AI chatbots have a promising application in higher education, including research and publication. There are growing studies on the opportunities, benefits, and challenges of chatbot usage in research. However, there is a lack of studies on the ethical concerns relating to chatbot usage in the scientific research process. This study addressed this matter and explored the ethical concerns associated with chatbot usage and strategies for mitigating these concerns. The research explored how academics and scientific research leaders made their ethical decisions for the use or non-use of chatbots in their scientific research process.

This research showed that ethical concerns are not only associated with legal, technical, and informational concerns but also have social concerns and implications for the sustainability of scientific research. The results identified ten major concerns, including privacy, confidentiality and data protection, bias and inaccuracy, accountability and responsibility, authorship, copyright and licenses, fairness and data accessibility, hallucinations, recency of data and information, validation and assessment of information, human skills and interaction, and the presence of

pseudoscience. Many of these concerns were also found in earlier studies on AI use in research and education (Chaka, 2023; Chaka, 2024; Popenici, 2023; van Dis et al., 2023; Xames & Shefa, 2023). However, the results of this research showed that these ten concerns were related to either legal, technical or social aspects. In addition, these concerns are interrelated and overlapped, albeit they all contribute to the presence of fake, junk and/or pseudoscience (see Figure 1).

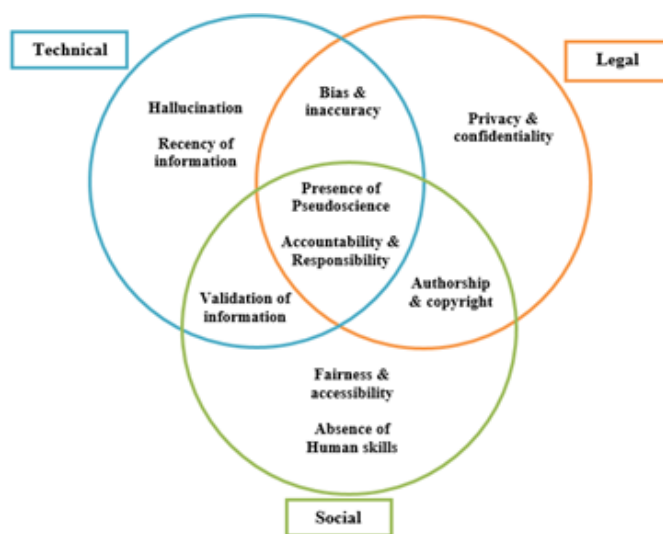


Figure 1. The overlap between ethical concerns associated with the use of chatbots.

Strategies for mitigating these concerns were suggested based on discussion with academics and scientific research leaders in Saudi Arabia, which is also of benefits to other countries. There was a consensus among all participants that a policy on the use of AI and AI chatbots in research should be developed as soon as possible to manage the irresponsible use of AI in research and publication. Aligning with previous research (Rudolph et al., 2023; Hasanien & Sobaih, 2023), this policy should manage the concerns related to privacy and data protection, authorship, copyright, and licenses. The policy should maintain the accountability and responsibilities of system creators and users. They should act ethically and responsibly in building the system, training data, and using the outputs. Academics have to be aware of chatbot limitations such as hallucination, data bias or inaccuracy, which require further validation and evaluation by them and experts in the field (Xames & Shefa, 2023). There is a need for accountability in developing the system and responsibility in using the outputs and considering the limitations of chatbots. It is important that academics adopt responsible use of AI chatbots and validate the outputs to prevent the presence of fake or junk science and ensure the sustainability of scientific research.

The ethical dilemma regarding the use of chatbots in research and publication arises when each group of participants (academics and leaders of scientific research) has its own justification and decisions about the use or non-use of chatbots in research. Most of the higher education institutions in Saudi Arabia force their academic staff, particularly international staff, to publish at least one international journal article annually. Hence, academics who did not have earlier experience in publication were looking

for assistance and found chatbots a great tool to facilitate this for them. The results confirmed that academics believe that chatbots have many advantages and few disadvantages. They were more interested in their advantages. However, leaders of scientific research believe it has a long-term impact and research integrity would be at risk. The first group considers the utilitarian ethical model that found the advantages of chatbot use in research and publication extremely exceed the disadvantages. This finding coincides with the utilitarianism approach, which was built on the theory of morality that supports activities, which foster positive outcomes and oppose those that generate harm or negative outcomes (Donagan, 1977). However, the second group members argued that these disadvantages put the integrity and sustainability of scientific research at risk. Hence, they believe in "virtue" or "common good" approaches to ensure positive long-term impact. Those who belong to virtue ethics are not concerned with codes of morality. However, they believe in the quality of consistently acting in harmony with their values (Crossan et al., 2013). In other words, those leaders are sacrificing morality for the "common good". Therefore, they were cautious about the use of AI chatbots without a clear policy. Hence, the policy of AI use in research and publication is urgently needed to address these concerns and make the best use of technology and AI in research and publication.

### Conclusion and future research directions

This research provides an insight on the ethical concerns and ethical dilemma relating to chatbot usage in research and publication. The research showed ten ethical concerns connected with legal, technical, and social concerns (Figure 1). The research showed that the ethical concerns are overlapped and are connected to irresponsible use of chatbots and their generated contents, which has an impact on the spread of pseudoscience. The research provided strategies for mitigating the risk associated with these concerns. The dilemma for the use or non-use of chatbots was related to how to view and use the chatbots. When academics found the chatbots' benefits overcame their challenges, they intended to use them intensively, employing a utilitarian approach. However, others believe that the disadvantages would have negative consequences, especially in the long term. They do not encourage their usage and believe in a "virtue" or a "common good" approach.

This research draws on a qualitative study with a sample of participants from public universities in Saudi Arabia who came from different backgrounds and cultures; this research did not analyze the participants' backgrounds or cultures and link them with their opinions. Further research could undertake a quantitative approach with a more significant sample to examine the variables that affect their ethical decision to use various chatbots in research and publication. A comparative study of the researchers' perceptions from various countries would be another avenue for future research.

**Acknowledgements:** This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi

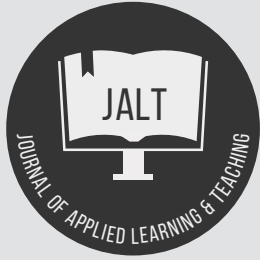
Arabia (GRANT5527).

### References

- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, 2(15), 100006. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Aiumtrakul, N., Thongprayoon, C., Suppadungsuk, S., Krisanapan, P., Miao, J., Qureshi, F., & Cheungpasitporn, W. (2023). Navigating the landscape of personalized medicine: The relevance of ChatGPT, BingChat, and Bard AI in nephrology literature searches. *Journal of Personalized Medicine*, 13(10), 1457. <https://doi.org/10.3390/jpm13101457>
- Al Lily, A. E., Ismail, A. F., Abunaser, F. M., Al-Lami, F., & Abdullatif, A. K. A. (2023). ChatGPT and the rise of semi-humans. *Humanities and Social Sciences Communications*, 10(1), 1–12. <https://doi.org/10.1057/s41599-023-02154-3>
- Al-Abdullatif, A. M. (2023). Modeling Students' perceptions of chatbots in learning: Integrating technology acceptance with the value-based adoption model. *Education Sciences*, 13(11), 1151. <https://doi.org/10.3390/educsci13111151>
- Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. In Ö. Aydın (Ed.), *Emerging computer technologies 2* (pp. 22-31). İzmir Akademi Derneği. <http://dx.doi.org/10.2139/ssrn.4308687>
- Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, 6(2), 94-104. <https://doi.org/10.37074/jalt.2023.6.2.12>
- Chaka, C. (2024). Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. *Journal of Applied Learning and Teaching*, 7(1), 1–12. <https://doi.org/10.37074/jalt.2024.7.1.14>
- Chen, T. J. (2023). ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association*, 86(4), 351-353. <https://doi.org/10.1097/JCMA.0000000000000900>
- Cheong, R. C. T., Unadkat, S., Mcneillis, V., Williamson, A., Joseph, J., Randhawa, P., Andrews, P., & Paleri, V. (2023). Artificial intelligence chatbots as sources of patient education material for obstructive sleep apnoea: ChatGPT versus Google Bard. *European Archives of Oto-Rhino-Laryngology*, 281, 985–993. <https://doi.org/10.1007/s00405-023-08319-9>

- Cilesiz, S. (2011). A phenomenological approach to experiences with technology: Current state, promise, and future directions for research. *Educational Technology Research and Development*, 59, 487–510. <https://doi.org/10.1007/s11423-010-9173-2>
- Colby, K. M., Hilf, F. D., Weber, S., & Kraemer, H. C. (1972). Turing-like indistinguishability tests for the validation of a computer simulation of paranoid processes. *Artificial Intelligence*, 3, 199–221. [https://doi.org/10.1016/0004-3702\(72\)90049-5](https://doi.org/10.1016/0004-3702(72)90049-5)
- Crossan, M., Mazutis, D., & Seijts, G. (2013). In search of virtue: The role of virtues, values and character strengths in ethical decision making. *Journal of Business Ethics*, 113, 567–581. <https://doi.org/10.1007/s10551-013-1680-8>
- Donagan, A. (1977). *The theory of morality*. University of Chicago Press, Chicago. <https://press.uchicago.edu/ucp/books/book/chicago/T/bo5965934.html>
- Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report*, 20(9), 1408–1416. <http://dx.doi.org/10.46743/2160-3715/2015.2281>
- Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Parlikad, A. K., Stankovski, V., Abraham, A., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Dustdar, S., Sakellariou, R., Uhlig, S., & Buyya, R. (2024). Transformative effects of ChatGPT on modern education: Emerging Era of AI Chatbots. *Internet of Things and Cyber-Physical Systems*, 4, 19–23. <https://doi.org/10.1016/j.iotcps.2023.06.002>
- Gonsalves, R. A. (2023). *Using ChatGPT as a creative writing partner - part 2: Music*. Towards Data Science. <https://towardsdatascience.com/using-chatgpt-as-a-creative-writing-partner-part-2-music-d2fd7501c268>
- Hasanein, A. M., & Sobaih, A. E. E. (2023). Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. *European Journal of Investigation in Health, Psychology and Education*, 13(11), 2599–2614. <https://doi.org/10.3390/ejihpe13110181>
- Heller, B., Proctor, M., Mah, D., Jewell, L., & Cheung, B. (2005, June). Freudbot: An investigation of chatbot technology in distance education. In *EdMedia+ innovate learning* (pp. 3913–3918). Association for the Advancement of Computing in Education (AACE). [https://www.researchgate.net/publication/242084006\\_Freudbot\\_An\\_Investigation\\_of\\_Chatbot\\_Technology\\_in\\_Distance\\_Education](https://www.researchgate.net/publication/242084006_Freudbot_An_Investigation_of_Chatbot_Technology_in_Distance_Education)
- Huang, D., Markovitch, D. G., & Stough, R. A. (2024). Can chatbot customer service match human service agents on customer satisfaction? An investigation in the role of trust. *Journal of Retailing and Consumer Services*, 76, 103600. <https://doi.org/10.1016/j.jretconser.2023.103600>
- Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIED) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching*, 6(2), 332–345. <https://doi.org/10.37074/jalt.2023.6.2.29>
- Iodice, G., Clemente, L., & Bifulco, F. (2024). The role of chatbot technology in the cultural value proposition: A managerial perspective in human-like interaction. In *The use of artificial intelligence in digital marketing: Competitive strategies and tactics* (pp. 166–187). IGI Global. <http://dx.doi.org/10.4018/978-1-6684-9324-3.ch007>
- Joseph, O. U., Arikpo, I. M., Victor, O. S., Chidirim, N., Mbua, A. P., Ify, U. M., & Diwa, O. B. (2024). Artificial Intelligence (AI) in academic research. A multi-group analysis of students' awareness and perceptions using gender and programme type. *Journal of Applied Learning and Teaching*, 7(1), 1–17. <https://doi.org/10.37074/jalt.2024.7.1.9>
- Kim, S. G. (2023). Using ChatGPT for language editing in scientific articles. *Maxillofacial Plastic and Reconstructive Surgery*, 45(1), 13. <https://doi.org/10.1186/s40902-023-00381-x>
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7), 5614. <https://doi.org/10.3390/su15075614>
- Le, X. C., & Nguyen, T. H. (2024). The effects of chatbot characteristics and customer experience on satisfaction and continuance intention toward banking chatbots: Data from Vietnam. *Data in Brief*, 52, 110025. <https://doi.org/10.1016/j.dib.2023.110025>
- Lee, J. Y. (2023). Can an artificial intelligence chatbot be the author of a scholarly article? *Journal of Educational Evaluation for Health Professions*, 20(6), 1–6. <https://doi.org/10.3352/jeehp.2023.20.6>
- Li, K. (2024). *Generative AI's wild 2023*. Reuters 2023 Year in Review. <https://www.reuters.com/world/year-in-review/>
- Liebrez, M., Schleifer, R., Buadze, A., Bhugra, D., & Smith, A. (2023). Generating scholarly content with ChatGPT: Ethical challenges for medical publishing. *The Lancet Digital Health*, 5(3), e105–e106. [https://doi.org/10.1016/S2589-7500\(23\)00019-5](https://doi.org/10.1016/S2589-7500(23)00019-5)
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- Lund, B. D., & Naheem, K. T. (2023). Can ChatGPT be an author? A study of artificial intelligence authorship policies in top academic journals. *Learned Publishing*, 37(1), 13–21. <https://doi.org/10.1002/leap.1582>
- Manigandan, L., & Sivakumar, A. (2024). Chatbot research: Unveiling evolutionary trends and collaborative pathways through bibliometric analysis. *Multidisciplinary Reviews*, 7(3), 2024045. <https://doi.org/10.31893/multirev.2024045>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education

- as explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>
- Neubauer, B. E., Witkop, C. T., & Varpio, L. (2019). How phenomenology can help us learn from the experiences of others. *Perspectives on Medical Education*, 8, 90–97. <https://doi.org/10.1007/s40037-019-0509-2>
- O'Connor, S., & ChatGPT. (2022). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice*, 66, 103537–103537. <https://doi.org/10.1016/j.nepr.2022.103537>
- OpenAI. (2022, December 19). *ChatGPT: Optimizing language models for dialogue*. <https://blog.cloudhq.net/openai-chatgpt-optimizing-language-models-for-dialogue/>
- Popenici, S. (2023). The critique of AI as a foundation for judicious use in higher education. *Journal of Applied Learning and Teaching*, 6(2), 378-384. <https://doi.org/10.37074/jalt.2023.6.2.4>
- Ram, B., & Pratima Verma, P. V. (2023). Artificial intelligence AI-based chatbot study of ChatGPT, Google AI Bard and Baidu AI. *World Journal of Advanced Engineering Technology and Sciences*, 8(01), 258-261. <https://doi.org/10.30574/wjaets.2023.8.1.0045>
- Rudolph, J., Tan, S., & Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care*, 27(75), 1-5. <https://doi.org/10.1186/s13054-023-04380-2>
- Stahl, B. C., & Eke, D. (2024). The ethics of ChatGPT—Exploring the ethical issues of an emerging technology. *International Journal of Information Management*, 74, 102700. <https://doi.org/10.1016/j.ijinfomgt.2023.102700>
- van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224-226. <https://doi.org/10.1038/d41586-023-00288-7>
- Waisberg, E., Ong, J., Masalkhi, M., Zaman, N., Sarker, P., Lee, A. G., & Tavakkoli, A. (2023). Google's AI chatbot "Bard": A side-by-side comparison with ChatGPT and its utilization in ophthalmology. *Eye*, 1-4. <https://doi.org/10.1038/s41433-023-02760-0>
- Wang, C., Liu, S., Yang, H., Guo, J., Wu, Y., & Liu, J. (2023). Ethical considerations of using ChatGPT in health care. *Journal of Medical Internet Research*, 25, e48009. <https://doi.org/10.2196/48009>
- Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>
- Xames, M. D., & Shefa, J. (2023). ChatGPT for research and publication: Opportunities and challenges. *Journal of Applied Learning & Teaching*, 6(1), 390-395. <https://doi.org/10.37074/jalt.2023.6.1.20>
- Yang, X., Gao, J., Xue, W., & Alexandersson, E. (2024a). *PLLaMa: An open-source large language model for plant science*. arXiv preprint arXiv:2401.01600.
- Yang, X., Wilson, S. D., & Petzold, L. (2024b). *Quokka: An open-source large language model chatbot for material science*. arXiv preprint arXiv:2401.01089.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Generative Artificial Intelligence in distance education: Transformations, challenges, and impact on academic integrity and student voice

---

Kershnee Sevnarayan<sup>A</sup>                      *A*                      *University of South Africa*

Mary-Anne Potter<sup>B</sup>                      *B*                      *University of South Africa*

---

### Keywords

---

Academic integrity;  
artificial intelligence (AI);  
ChatGPT;  
distance education;  
generative artificial intelligence (GenAI);  
student voices.

### Abstract

---

Generative artificial intelligence (GenAI) has reshaped distance education by prompting a shift towards student-centred initiatives to promote responsible AI usage. This study explores the transformative impact of GenAI in distance learning and focuses on academic integrity and student voices. This study uses the technology acceptance model to investigate how GenAI influences distance education. Three objectives guide the study: (1) exploring the transformative effects of GenAI in distance education, (2) understanding its impact on academic integrity, and (3) examining its influence on students' academic voices in a South African open distance and e-learning university. Qualitative data was gathered through interviews with lecturers, open-ended evaluation questions with administrative staff, and focus group discussions with first-year students in an academic writing module. Findings highlight the need to bridge the gap between negative perceptions of AI's impact on academic integrity and positive views on its potential to boost student confidence in learning. This research study aims to analyse GenAI's role in distance education and provide insight into its potential, challenges, and strategies to ensure academic integrity and preserve students' voices.

---

### Correspondence

esevark@unisa.ac.za <sup>A</sup>

---

### Article Info

Received 18 March 2024  
Received in revised form 11 April 2024  
Accepted 29 April 2024  
Available online 29 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.41>



## Introduction

Generative artificial intelligence (GenAI) stands as a pivotal force in higher education institutions (HEIs) (Chugh et al., 2023; Maphoto et al., 2024; Ogata et al., 2024; Qasim et al., 2022) and serves as a foundation for innovation that not only reshapes traditional paradigms but also brings in an era of unprecedented possibilities for teaching and learning (Baidoo-Anu & Ansah, 2023; Fullan et al., 2023; Rudolph et al., 2024). Amidst this paradigm shift, one facet stands out – the introduction of GenAI, such as ChatGPT. Far from being a mere technological augmentation, the integration of ChatGPT into distance education prompts a re-evaluation of pedagogical norms, raising critical questions about its transformative influence on assessment methodologies (Naidu & Sevnarayan, 2023), academic integrity (Sevnarayan & Maphoto, 2024), and the dynamics between student voices (Ali et al., 2023) and authorship. As education embraces the capabilities of GenAI, it is essential to acknowledge the positive disruptions it introduces. Amid the potential for a paradigm shift in virtual interactions through adaptive responses (Rasul et al., 2023), there lies a need for careful exploration to consider the challenges identified by Cotton et al. (2023) and Naidu and Sevnarayan (2023). In the context of AI tools such as Copy.AI, QuillBot, Grammarly, Jasper AI, Notion AI, Jenni.AI, Wordtune, ProWritingAid, AI Writer, Rytr, and Writesonic, which collectively display diverse capabilities (Ladha et al., 2023), we explore specifically ChatGPT's capacity to transform online assessment methodologies. This exploration aims to empower lecturers to tailor learning experiences to individual needs.

Scholars have argued that the “proliferation of GenAI technologies like ChatGPT poses significant challenges to traditional assessment methods, particularly essays and online examinations” (Rudolph et al., 2024, p. 11). Consequently, Popenici (2023) offers a critical view of AI and states that “[a]ny technological solution and adoption involve a certain ideological choice and influence, consciously or not” (p. 381). This statement emphasises the need to analyse the ideological implications of AI integration in education critically, and it suggests that it may not always align with the values and aims of educational institutions. Popenici (2023) further emphasises, “The blind trust in and adoption of new tech by educators... becomes even more dangerous in the era of AI. The challenge ahead for education is to become users of AI for the benefit of our students and institutions rather than simple subjects of AI” (p. 381). This stance highlights the importance of approaching AI integration in education with vigilance and critical inquiry rather than unbridled enthusiasm. Popenici et al. (2023) note that “the advantage of ChatGPT is that it came as a mirror for education. It shows where we are because we are completely unprepared for reality. We lose perspective on what matters when we have this engine of mediocre text” (p. 323). This statement reveals educational institutions' unpreparedness for AI technology's realities. When our students generate text that may be perceived as mediocre or lacking in depth, ChatGPT exposes the limitations of current educational approaches in encouraging critical thinking, creativity, academic integrity and authentic student voices. They go on further to add that:

The very amusing thing, and it is laughable, is that technology is showing us how far we are from what we should do. The risk is that we are going to lose our legitimacy entirely. It's a massive challenge because we turned assessment into this industrial process of mass assessment, with no quality, no look of originality, and need of substance. This is what you have to submit; use citations; use good grammar, and good syntax, and you don't do massive mistakes. It's good to go. You graduate. You're good. It's fantastic when we turn the whole system to this; it's just that we lost the plot, and then it is a disgrace. Technology is showing us how much we are at risk. It is striking at the core of education. This is a consequential moment (Popenici et al., 2023, p. 323).

Similarly, Lindgren (2023) warns against the naturalisation of dominant views and priorities within AI discourse and argues the importance of critical analysis in questioning the social and political implications of innovation, progress, control, and efficiency. Concurrently, we confront concerns about the authenticity of students' work, academic integrity, and the impact on students' intellectual autonomy. Popenici et al. (2023) argue, “If you reduce learning to assessment and the assessment can be outsourced by students to just write a sentence and think a bit about the text, you have no motivation. Why would I do that? Why would I learn anything? Because I can just give it this AI solution. The kind of implications for universities are massive” (p. 324). The integration of GenAI in HEIs represents not just a technological leap but a fundamental shift in the distance educational context.

According to Yu (2024), the integration of ChatGPT into distance contexts transcends mere technological augmentation; it marks a profound reconfiguration of the higher education trajectory. The foundation of remote learning and online assessment undergoes a metamorphosis, challenging conventional wisdom about evaluation. The reliability of AI-generated responses to reflect a student's comprehension and originality becomes a central question and demands a reconsideration of assessment methodologies in this new paradigm. For Perkins (2023), academic integrity, a foundation of educational excellence, faces a formidable test in the era of ChatGPT. The potential for blurred lines between collaboration and unauthorised assistance looms large and raises concerns about plagiarism and academic dishonesty. Does the integration of GenAI empower students to articulate their thoughts more effectively, or does it dilute the essence of authorship, posing a threat to the authenticity of their academic contributions? In this article, we explore the challenges and opportunities presented by AI and provide insights into how lecturers and HEIs can engage with these changes effectively. We pose the following research questions in this article:

- How has ChatGPT influenced and transformed distance education?
- How does ChatGPT impact on academic integrity in distance education?

- In what ways does ChatGPT affect students' academic voice in the context of distance learning?

## Literature review

We acknowledge the scope and contribution of literature in advancing the debate around the challenges and opportunities GenAI presents within the higher education context. However, preference is given to the most pertinent and contemporary peer-reviewed scholarship on GenAI's and ChatGPT's specific impact on tertiary teaching and learning.

### Distance education in the age of AI and ChatGPT

Since its November 2022 launch (OpenAI, 2022), OpenAI's ChatGPT-3 has gained attention, amassing over 100 million users by January 2023. Within HEIs, concerns about GenAI, driven by profit motives, have been voiced (Ormond, 2023). However, Mollick and Mollick (2022) counter this view by proposing that GenAI can enhance learning by overcoming barriers like improving transfer, debunking explanatory depth illusions, and training critical evaluation skills. Our study focuses on the latter aspect as students' abilities to critically evaluate information shapes academic voice and integrity. Distance education, amid emerging GenAI, presents unique opportunities and challenges. While GenAI can personalise learning, challenges include bias, overreliance on AI hindering critical thinking, and access disparities (Bozkurt & Sharma, 2023). Bozkurt and Sharma (2023) advocate for personalised, adaptive, student-centred approaches that are feasible in the Global North but challenging in the less technologically developed Global South. In addition, open distance e-learning (ODEL) institutions face 'transactional distance' (Moore, 2013), separating lecturers from students. With students managing their learning, support is mediated through technology like learner management systems and communication platforms (Bozkurt & Sharma, 2023). In the student-centred ODEL approach, ChatGPT would form part of such mediatory technology.

Research by Holmes and Porayska-Pomsta (2023) and Maphoto et al. (2024) reveal a gap in understanding the impact of GenAI on academic integrity in the Global South. This relates specifically to the potential of South African ODEL HEIs to facilitate a dialogue that extends beyond student-facilitator interactions to include conversations initiated with GenAI systems (Bozkurt & Sharma, 2023). However, caution is advised, as GenAI should complement, not replace, personalised support from lecturers (Bozkurt & Sharma, 2023). ChatGPT, as a secondary digital facilitator, automates traditional tasks, reducing transactional distance but necessitating continued lecturer guidance. Bozkurt and Sharma emphasise the integration of GenAI with core educational values such as equity, diversity, and inclusivity and consider the dynamic nature of ODEL contexts (2023, p. vi). In this context, the widespread use of distance education and the rise of GenAI highlight the need for thoughtful integration that aligns with educational values in the digital age.

### Distance education, academic integrity, and ChatGPT

Debates surrounding the integration of GenAI, particularly ChatGPT-3, in education, especially distance education, centre on ethical concerns. Oppenheimer (2023) argues that while AI systems may increase access to information, concerns about academic integrity existed before, and AI does not fundamentally alter these dynamics. Tlili et al. (2023) propose an ethical and interactive integration of GenAI systems to enhance the development of twenty-first-century competencies. Identifying these competencies stems from addressing what Weinberger (2007) terms 'a new digital disorder'. Dede (2009) sees the development of twenty-first-century skills as a response to 'disorderly' knowledge co-creation and sharing. As technology disrupts learning, especially in the post-COVID world, a holistic appraisal of stakeholders' experiences in teaching and learning is necessary in the ODEL context. The impact of technology disruption is evident during learning assessments, particularly in plagiarism. Jones and Sheridan (2015) note that plagiarism affects student equity and diminishes qualities aligned with academic voice. The incorporation of GenAI technologies like ChatGPT in student resources necessitates a re-evaluation of assessment strategies. Oppenheimer (2023) suggests focusing on integrity training, cultivating a healthy campus culture, and reducing incentives to cheat as effective strategies against academic dishonesty. However, ChatGPT's reliance on various sources and students' ability to discern source credibility raise concerns about its reliability as an educational resource. OpenAI's President, Brockman (2023), acknowledges the system's "emergent (unanticipated) capabilities," making it volatile in its current form. As debates persist, careful consideration of the ethical integration of GenAI, assessment strategies, and the impact on academic integrity in distance education is crucial. Sullivan et al. (2023) have further directed such debates toward how developing policies related to the use of GenAI would take time and be directly influenced by a longitudinal evaluation of how such technology is used or misused (p. 35).

### Generative AI's influence on academic voice in distance education

As indicated above, academic dishonesty impacts the ability to develop an academic voice. McQuillan (2021) describes voice as being founded on "original ideas and thoughts and [that] it is used to establish credibility" within the academic context (p. 32). She continues to emphasise that originality in thinking is fundamental to developing an academic voice (McQuillan, 2021). McQuillan, therefore, equates academic voice with originality and responsibility—a connection also noted by Thompson (2012, p. 121). In terms of the concerns regarding AI's potential challenge to originality and responsibility, Holmes et al. (2023) highlight four key concerns: information bias, AI's capacity to make autonomous decisions, which could impact outcomes quite severely, privacy in the exchange of personal data, and the "potential [for AI] to be used for malicious purposes" (p. 97). Concerning the development of academic voice among undergraduate students, exposure to information mediated by GenAI systems such as ChatGPT means that students need to be taught better and more efficient ways to discern

between credible and non-credible information. Therefore, stakeholders in the learning process within the ODeL context cannot deny that an additional nonhuman stakeholder has emerged. In response, Holmes et al. (2023) propose a new stakeholder framework that acknowledges AI as centrally influencing the interaction between students, teachers, and institutions (p. 101). However, they are cautious in asserting how GenAI's influence—its algorithms and programming—will be enacted in the future (Holmes et al., 2023).

This framework is relevant to the ODeL context in that much of the interaction between teachers, students, and institutions is mediated by technology. Mediation calls into question the capacity of students to develop an authentic academic voice through AI systems like ChatGPT or whether that voice is vulnerable to manipulation. The higher instances of plagiarism among students suggest that the development of an authentic academic voice is under threat. However, Rudolph et al. (2023) assert that the perceived threat may be because of a "bureaucratic inertia" that is struggling to adapt to the capacity of students to engage with GenAI (p. 354). Rudolph et al. (2023) further recommend that GenAI has the capacity to enhance the facilitation of students' curiosity, experimentation, and collaboration as a learning outcome. However, traditional assessment policies do not fully embrace this measure of student learning. Ifelebuegu (2023) comments on the debate between authentic and AI-assisted assessment, which is more complicated in online learning contexts. Student voice, as implied by Ifelebuegu (2023), is threatened when the dialogic potential between students and GenAI tools is not employed to enhance critical thinking:

The advent of AI chatbots has introduced a unique challenge to the integrity of online assessments, leading educators to reevaluate traditional assessment methods. As we navigate this landscape, it is clear that assessments must evolve to maintain their authenticity and effectiveness in promoting meaningful learning. This exploration has underscored the importance of reshaping assessments to value higher-order cognitive skills, problem-solving, creativity, and collaborative abilities. Authentic assessments such as open-ended tasks, project-based assignments, collaborative assessments, and portfolio-based assessments not only align with these values but also pose a significant challenge for AI chatbots to replicate or assist in, thereby preserving their integrity. AI may also aid assessment rather than just being a danger (Ifelebuegu, 2023, p. 389).

From a student-centred perspective, Chan and Hu (2023) note the challenge that AI presents to the development of holistic competencies, such as the capacity for creative and critical thinking, which inform academic voice (p. 11). Baker and Smith (2019) contextualise the incorporation of GenAI in education as student-centred and further assert the need to consider two other key stakeholders in evaluating how GenAI is incorporated: teachers and systems or administrators. Ahmad et al. (2024) noted that each stakeholder has differing views of the benefits and threats presented through GenAI. While Ifelebuegu et al. (2023) note

the benefits that include assisting with academic writing, facilitating personalised learning, and interdisciplinary education, they also highlight the threats of information bias and misinformation, plagiarism, and an overreliance on technology at the expense of critical thinking development in students.

### Using the technology acceptance model for AI

The technology acceptance model (TAM), developed by Davis (1989; 1993), serves as a lens to explore the factors influencing the acceptance and usage of technology. In this context, external factors represent the features of ChatGPT, assessing its adaptability to personalised learning experiences and its potential to reshape assessment methodologies in distance education (Davis, 1989). The study explores cognitive responses (Davis, 1989), mainly the perceived ease of use and usefulness of ChatGPT. It examines the effortlessness of interaction and how much it enhances academic experience in a distance education context. The affective response (Davis, 1993), captured by attitudes toward using technology or behavioural intentions, highlights how students and lecturers perceive and intend to utilise ChatGPT. The outcome reflects the active integration of ChatGPT into academic activities (Davis, 1989). TAM outlines a three-stage process for technology acceptance, as we illustrate in Figure 1 below.

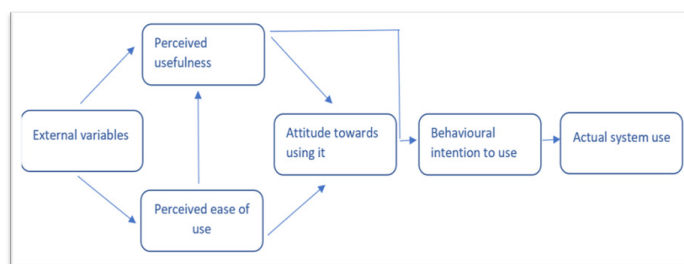


Figure 1. The technology acceptance framework (adapted from Davis, 1989).

External factors trigger cognitive responses, such as perceived ease of use and usefulness, leading to an affective response and influencing user behaviour. The model predicts behaviour based on perceived ease of use, perceived usefulness, and behavioural intention. Affective responses, particularly attitudes toward behaviour, play a crucial role in determining the likelihood of technology adoption. While perceived usefulness directly affects one's use behaviour, perceived ease of use indirectly influences it by supporting the effect of usefulness. If an application is perceived as easy to use, it is more likely to be considered useful, stimulating technology acceptance (Davis, 1989, 1993). Applying TAM to the study provides a structured framework for analysing ChatGPT's influence on distance education (Davis, 1989). It enables an examination of how external factors impact cognitive responses and influence the acceptance and integration of ChatGPT into the context under study (Davis, 1989). The model is integral in addressing concerns about academic integrity by assessing how perceived ease of use and usefulness influence attitudes and intentions, thus contributing to a critical understanding of the ethical considerations surrounding ChatGPT's application in

education (Davis, 1989). Furthermore, TAM facilitates an exploration of the affective responses and reveals insights into how students perceive ChatGPT's role in shaping their academic voice and originality.

## Methodology

### Research approach and design

This study employed a qualitative case study research approach (Baskarada, 2014) to explore the transformative influence of ChatGPT in distance education. A case study design was deemed appropriate for its ability to provide an in-depth understanding of the phenomena within its real-life context. The study focused on the Department of English Studies at an ODeL university in South Africa.

### Research context and population

The research was targeted at first-year students enrolled in an academic writing module. The decision to focus on first-year students stems from their diverse backgrounds that reflect a wide range of linguistic, social, and cultural perspectives. However, challenges, such as limited access to technology, highlight the need for tailored support. Moreover, a digital divide among lecturers highlights the complexity of addressing these issues within the department. This module reflects a diverse student body with varying linguistic, social, and cultural backgrounds. Notably, many students faced challenges such as limited access to technological tools and diverse personal responsibilities, including part-time or full-time work and caregiving responsibilities. The lecturers who teach first-year students in the department, spanning various age groups, indicate a digital divide within the teaching staff.

### Population, sampling, and research instruments

The population for this study comprised the entire cohort of 14,000 first-year students enrolled in the Academic Writing module under study. The sampling strategy was purposive, targeting ten first-year lecturers for one-on-one semi-structured interviews to address the first research question. In addition, five administrative staff members were selected to respond to the second research question through open-ended email evaluation questions. For the third research question, 20 students who spoke English as an additional language (EAL) were invited to participate in a Microsoft Teams focus group discussion (FGD).

### Data collection

The study was conducted during the second semester of 2023, from August to December 2023. Ten lecturers were interviewed individually via e-mail, five administrative staff responded to open-ended email evaluation questions over two weeks, and a Microsoft Teams FGD involved 12 out of the 20 invited students, which lasted approximately 1.5 hours. Thematic analysis was used to analyse the collected data,

which were organised according to the following themes:

- How ChatGPT technology influenced and transformed distance education.
- The impact of ChatGPT on academic integrity in distance education.
- How ChatGPT affects students' voices in distance education.

### Ethical considerations

Informed consent was obtained from all participants, ensuring their voluntary participation. Participants were assured of confidentiality, and the module name, the name of the university, and all participants were anonymised in the presentation of findings to protect their identities. Lecturers in the study are called L1, students S1, etc., and administrative officers, A1 and so on. The study adhered to ethical guidelines, and approval was obtained from the relevant institutional ethics review board at the university under study. Moreover, participants were informed of their right to withdraw from the study at any point without consequences.

## Findings and discussion

### How ChatGPT technology influenced and transformed distance education

The responses from lecturers regarding the impact of ChatGPT were limited, with only two of the targeted ten lecturers providing critical insights in this regard. Both responses were sceptical of the effects of ChatGPT on and/or transformed distance education.

### *Perceived negative impact on academic integrity*

With literature noting the impact of academic dishonesty on voice and the ability to think critically and creatively (Jones & Sheridan, 2015), both lecturers emphasised the negative impact of ChatGPT on academic integrity. L1 responded to the open-ended evaluation question by stating, "Unfortunately, in my experience, it has made academic dishonesty more rife". L1 drew on their experience in assessing student work to provide examples of the impact of ChatGPT on student performance:

We have found students to not only have fed our questions into it and submitted the answers it provided (to varying levels of soundness; some of the AI answers are decent, some are nonsensical and only barely relevant to the question), but we have also found students to have shared these answers amongst each other in the distance-based context, with the result being that a not-insubstantial number of students submitted the exact same AI-generated answers.

This view is echoed in L2's response, where they observe that "ChatGPT perhaps intensifies what Google and the internet have long established: the sense that knowledge and understanding are immediately available and accessible through the provision of the appropriate search terms or application of the correct technique", thereby affecting critical and creative thinking. Authentic assessment as measuring independent, critical thinking is, therefore, compromised according to lecturers' perspectives, an assertion supported by Bozkurt and Sharma (2023). However, reliable and actionable strategies to mitigate this are not indicated. The responses offered by lecturers contradict TAM in that positive acceptance is countered by intense skepticism and heightened caution. In reviewing the two responses, they are contextualised within the scope of older, more established systems of assessment encountering new disruptive forms of information access. This is echoed in the concerns Chan and Hu (2023) expressed regarding the threat generative AI poses to the development of holistic critical competencies.

Teaching and learning transformation within ODeL contexts Though L1 did not acknowledge the need for novel approaches to teaching and learning, and in particular, assessment, L2 was more open to considering how ChatGPT's introduction into the teaching and learning context requires a shift in strategy. They observe the following:

My sense is that it's transforming, rather than fully transformed, so I don't think I can provide a definitive answer. It has shifted my sense of how I should approach my teaching. I feel a heightened need to guide students to the sense that they themselves are vital to what we call under blanket terms 'learning,' 'scholarship,' 'research', etc. – to help them develop a sense that the point of the exercise is a transformation in the learner, not the retrieval or repetition of what is already known.

L2 is introspective in terms of demonstrating awareness that the shift that is required in response to the disruption of GenAI cannot be levied on the student but should be guided by new teaching and learning practices, including assessment practices: "I'm experiencing serious doubts about assessment practices in this context—and even about the very principles which underlie assessment. This is a 'crisis' (a moment of decision?) not unique to the distance learning environment." The lecturer demonstrates the space that ODeL now occupies in terms of its response to disruption—in this case, the advent of GenAI—and the need to adopt new thinking practices.

In this sense, real-world concerns about GenAI's negative impact on existing teaching and learning practices encounter a positive call to transform teaching and learning practices to future-proof higher education in ODeL contexts. However, the general view is one of despondency in the face of a new challenge, making transformation seem insurmountable.

## **The impact of ChatGPT on academic integrity in distance education**

The responses from the five administrative staff members varied, which reflected a spectrum of perspectives on the impact of ChatGPT on academic integrity in distance education.

### ***Perceived positive impact on skills development***

An administrator emphasised ChatGPT's potential to contribute to critical thinking and independent research skills and portrayed it as a facilitator rather than a replacement for essential academic skills. A1 noted that "ChatGPT can act as a tool to promote critical thinking and independent research skills. I think it challenges students to use information critically, which makes learning more exciting." The emphasis on promoting critical thinking and independent research skills suggests that A1 sees ChatGPT as a facilitator rather than a replacement for these essential academic skills. The findings do not corroborate with scholars such as Rudolph et al. (2024) and Popenici et al. (2023), who caution against the blind adoption of new technologies in education and highlight the ideological implications and challenges posed by the integration of AI. While Rudolph et al. (2024) and Popenici et al. (2023) rightly emphasise the ideological impact and challenges of adopting AI, it is essential to recognise that the area of educational technology is multifaceted, and its impacts can vary significantly depending on the context, implementation, and the specific technology in question. The findings, however, resonate with the idea that AI technologies can supplement human capabilities and provide additional resources and support rather than diminishing the role of students' independent thinking (Bozkurt & Sharma, 2023). This aligns with TAM, where positive perceptions of a technology's usefulness and ease of use influence its acceptance. This perception of usefulness can shape positive attitudes and behaviours toward integrating ChatGPT in educational contexts. A1's positive outlook aligns with the literature suggesting that AI technologies, including ChatGPT, can enhance critical thinking and independent research skills (Mollick & Mollick, 2022).

### ***Concerns about authenticity and originality***

In contrast to A1, A2 expressed reservations that echo concerns in the literature about potential overreliance on AI-generated content challenging the authenticity and originality of students' work. He argues,

The risk of students relying too heavily on AI-generated content is concerning. It challenges the authenticity of their work and raises questions about the originality of their ideas. We have seen such instances this year where plagiarism was rife in assignments. This then affects us all, from lecturers to the students.

The acknowledgement that there have been instances of plagiarism raises immediate red flags about the tool's impact on the authenticity and originality of students'

work. A2 argued that it is concerning that students use AI to plagiarise in their assessments. This calls into question students' originality, their learning, and the integrity of the university. This aligns with existing literature highlighting AI technologies' challenges to academic integrity and the need for a vigilant approach (Jones & Sheridan, 2015; Maphoto et al., 2024; Oppenheimer, 2023). In addition, Lindgren (2023) stresses the importance of critical analysis in questioning AI innovation's social and political implications, particularly in education. These critical perspectives challenge the notion that ChatGPT facilitates skills development and urges lecturers and institutions to consider the broader ethical implications and societal consequences of AI integration in education. Moreover, Chaka (2024) argues that contemporary AI detectors and traditional anti-plagiarism tools should be combined with human reviewers and raters to differentiate between AI-generated and human-generated text. This aligns with broader concerns in the literature about the potential disruption of traditional learning paradigms, such as overreliance on AI hindering critical thinking, bias and access disparities, and the need for comprehensive strategies to mitigate risks (Bozkurt & Sharma, 2023). This finding also calls for lecturers and stakeholders in HEIs to educate students on the responsible and ethical use of AI. If students are not taught how to use AI responsibly and ethically, we must expect plagiarism and unethical AI usage from students who are not confident with writing using their voices.

### ***Importance of safeguards and transparency***

Administrators introduced a critical perspective and emphasised the need for assessment safeguards and proactive university guidance regarding ChatGPT's use. A3 suggested:

The university should guide us on how to deal with AI when it comes to assessments. There are talks of lecturers using it in assessments, but if this is to happen, we must ensure that students use it wisely. But I think the university needs to take a stand.

A3 suggested that the university should proactively guide lecturers on how to incorporate GenAI in assessments. While there are discussions about allowing lecturers to use GenAI in assessments, A3 highlighted the need to establish clear guidelines. Like A2's perspective, A3 raised concerns about the potential misuse or overreliance on AI-generated content but also suggested a focus on ensuring a smooth integration process to promote the acceptance and effective use of ChatGPT (Davis, 1989; Venkatesh & Davis, 2000). The perceived ease of use of ChatGPT, as highlighted by A1 and A3, may contribute to a positive attitude among administrative staff members. This ease of use can impact their willingness to adopt ChatGPT in academic assessments. The emphasis is ensuring that students view ChatGPT as a supplementary tool rather than a replacement for their ideas. This aligns with the call for faculty development programmes and workshops to guide lecturers in effectively integrating AI technologies into assessments and ensure that students use them as supplements rather than replacements for their

ideas (Bozkurt & Sharma, 2023). A4 and A5 further argued for the importance of transparency in using ChatGPT for assessments. They noted that students should be educated about the role of ChatGPT, its limitations, and the significance of their independent contribution to academic work. A5 argues, "Transparency is key. Students should be taught about the role of AI in assessments". This approach aims to create awareness of the ethical use of ChatGPT in HEIs and a balanced understanding of the technology's role in the learning process. This corroborates with TAM's emphasis on clear communication and education influencing technology acceptance (Davis, 1989). The stress on transparency also resonates with the principles of integrity and responsible AI usage (Holmes et al., 2023).

### **How ChatGPT affects students' voices in distance education**

#### ***Changing perspectives on student voice and AI integration***

Integrating ChatGPT into distance education introduces a transformative force beyond technological augmentation (Chugh et al., 2023). This change in thinking prompts a re-evaluation of pedagogical norms, particularly in terms of its impact on student voices and authorship (Ali et al., 2023). According to one student: "It's already tough because my home language is isiXhosa. I don't think a computer can understand what I'm trying to say better than a person" (S1). S1's hesitation about relying on AI due to language barriers and concerns about the computer understanding their thoughts reflects an initial skepticism that aligns with the broader literature on the apprehension towards AI in education (Chugh et al., 2023; Popenici et al., 2023; Rudolph et al., 2024). This skepticism is crucial to perceived ease of use, as students might find it challenging to utilise AI tools effectively. S1's concerns highlight the importance of addressing the ease with which students can interact with ChatGPT and emphasise the need for user-friendly interfaces and training. S2 expresses concerns about the potential impact of ChatGPT on their writing style and individual voice. "It might be a shortcut, you know? What if it...makes me sound less like me?" (S2). The fear of losing personal identity in their writing and the possibility of negative assessment consequences reflect worries about maintaining authenticity and uniqueness in academic work. S2's concern about ChatGPT as a 'shortcut' reflects perceived complexity (perceived ease of use). Their worry about the tool changing their writing style emphasises consideration of usefulness (perceived usefulness). The fear of sounding 'less like me' and potential assessment failure indicate emotional aspects influencing attitudes and behavioural intentions, aligning with TAM principles.

#### ***Student voice enhancement through AI***

As ChatGPT promises to transform online assessment methodologies (Naidu & Sevnarayan, 2023), the second subtheme explores students' perspectives of AI as a supportive tool. However, concerns about potential complacency in students' efforts necessitate a careful examination of the role of AI in education.

If AI can help me catch mistakes and suggest improvements, I'm all for it. It could save me a lot of time and stress, especially when deadlines are looming (S6).

I worry that relying too much on AI might make me lazy. I want to make sure that I do the work for me? (S11).

The contrasting views among students highlight the need to assess the implications of AI integration critically. S6's perspective aligns with the concept of perceived usefulness in TAM, where AI is viewed as a time-saving tool for error detection and improvement suggestions (Davis, 1989). This positive outlook resonates with the literature on the transformative potential of AI in enhancing efficiency and task performance in education (Rasul et al., 2023). On the other hand, S11's concern about overreliance on AI and potential complacency is consistent with discussions on the ethical dimensions of AI integration in education (Oppenheimer, 2023; Popenici et al., 2023; Rudolph et al., 2024). This resonates with the technology acceptance model (TAM). This model illustrated how attitudes and potential behavioural intentions are critical in technology acceptance (Davis, 1993). The apprehension expressed by S11 aligns with studies that emphasise the importance of responsible implementation to mitigate concerns related to dependency on AI (Cotton et al., 2023). With its emphasis on perceived ease of use and usefulness, TAM provides a critical understanding of students' attitudes toward AI and emphasises the need to consider affective responses and intentions. The notion of 'perceived ease of use' may be influenced by factors such as prior experience with technology, digital literacy, and socio-economic status. It is pertinent to acknowledge that students' readiness to accept and incorporate ChatGPT into their academic activities is significantly influenced by the perceived ease of use (Davis, 1989). Furthermore, the focus on individual attitudes and intentions in TAM overlooks the structural and systemic factors that shape the adoption and implementation of AI in education. A more critical approach would consider the political economy of AI, the role of neoliberalism in shaping education policy, and the impact of AI on the labour market and the future of work.

### ***Striking a balance between authenticity and learning***

The third subtheme explores questions about the impact on academic integrity and points to the need for ethical considerations in the integration of ChatGPT.

I'm worried that if I use AI too much, it might feel like I'm not really learning. I want to know I passed on my own strength (S7).

How do I know if the suggestions from AI make my writing better or just more 'correct' in a technical sense? (S8).

S7's apprehension about overreliance on AI speaks to the balance needed between technological support and preserving learning experiences. This concern, reflected

in broader discussions on the challenges of maintaining authenticity in AI-assisted education (Holmes et al., 2023), aligns with TAM's concept of perceived usefulness. The concern centres around the potential impact on agency in the learning process and emphasises the need to carefully integrate AI tools to enhance learning outcomes while preserving students' sense of accomplishment and reducing transactional distance (Davis, 1993; Moore, 2013). However, it is crucial to critically examine the power dynamics at play, as the increasing reliance on AI may perpetuate existing inequalities and reinforce dominant epistemologies. Moreover, emphasising efficiency and productivity in AI-driven learning environments may lead to a narrow focus on technical skills, neglecting critical thinking and creativity. Similarly, S8's inquiry into the effectiveness of AI suggestions in improving writing quality centres on the authenticity of the learning experience. This corroborates with literature highlighting the evaluation of AI-generated content and its compatibility with individual writing styles (Perkins, 2023). However, it is essential to interrogate the notion of authenticity in the context of AI-assisted learning and consider the potential for GenAI to both enable and constrain student agency. The literature foregrounds the potential benefits of AI in reshaping traditional paradigms (Baidoo-Anu & Ansah, 2023) and personalising learning experiences (Bozkurt & Sharma, 2023). However, the concerns raised by students align with the ethical considerations highlighted in the literature, such as the potential impact on academic integrity (Perkins, 2023) and the need for discernment in utilising AI resources (Holmes & Porayska-Pomsta, 2023). These concerns are technical and political, requiring a critical understanding of the intersections between technology, power, and pedagogy.

### **Limitations**

While this study provides insight into the transformative influence of ChatGPT in distance education, several limitations should be considered. Firstly, the focus on first-year students in a single department at an ODeL university in South Africa may need to be revised to include the generalisability of findings to other educational contexts. In addition, the reliance on qualitative methods, such as interviews and focus group discussions, may introduce bias or subjectivity in data collection and interpretation. Finally, purposive sampling may result in a non-representative sample and potentially overlook diverse perspectives from participants.

### **Conclusion and recommendations**

Within the scope of our study, several key findings emerged. Lecturers tended to be more pessimistic and cautious about the impact of ChatGPT on academic integrity. Unlike lecturers, administrative staff and students adopted a more transformative view of ChatGPT's potential in enabling learning, though it requires careful management. In filtering all responses through the TAM lens and orientating them within the ODeL context, it is evident that stakeholder attitudes towards the impact of ChatGPT on academic integrity and academic voice are not fully aligned in purpose. This has

created barriers to acceptance because of the prescriptive nature of the current teaching and learning context of ODeL institutions. However, such institutions present the greatest potential for transformation because of their orientation toward digital teaching and learning engagement. This may mean open conversations between stakeholders about academic integrity, which was flagged as a stakeholder concern to varying degrees. The discrepancy between the skepticism towards ChatGPT's ability to negatively impact academic integrity and the positive attitudes towards its ability to encourage more confident student interaction with learning materials compromises teaching and learning as authentic to the context within which it is enacted. This discrepancy feeds into an already-existing debate between authentic and AI-assisted assessment (Ifelebuegu, 2023). The emphasis on traditional assessment methods not only undermines student voice but also perpetuates a culture of standardisation, which is antithetical to the principles of authentic assessment. If lecturers rethink assessment strategies, they could create opportunities for students to engage in meaningful, self-directed and/or collaborative learning that values their unique voices and diverse sociocultural perspectives. The artificiality of GenAI is disrupting the academic values of integrity and voice, highlighting the inadequacy of traditional assessment to facilitate new forms of dialogic and collaborative learning.

The findings in this study demonstrate an urgent need for HEIs to rethink their approaches to teaching and learning in the age of GenAI. To encourage the power of GenAI in enabling students to utilise their voices, ODeL institutions should encourage open conversations, integrate GenAI into teaching and learning, and guide students with the technology towards building a community of practice. Lecturers can implement GenAI into their teaching through student-led discussions, reflective journals, peer review, critique, and student-generated questions, which promote ownership and exploration of topics. AI should not be used to elicit fearmongering in education but can be used as a tool to help students use their voices to generate prompts, generate ideas, improve their writing, edit their submissions and provide constructive critical feedback on peer work. In addition, lecturers can create multimodal assignments, allowing students to express themselves in diverse formats that reflect critical and creative thinking. Lecturers should create an inclusive and interactive space where students feel comfortable sharing their thoughts and ideas and develop critical thinking and communication skills even further. This is especially important with GenAI, as it enables students to utilise the potential of AI tools to support their learning while maintaining their unique voices and agency as distinct from AI. Most importantly, lecturers should be trained to teach students responsible and ethical use of GenAI and create a culture of digital literacy and accountability. In doing so, we can equip students with the skills to benefit from AI's potential while mitigating its risks and ensuring they become informed and ethical users of these powerful technologies. However, this cannot be fully realised without the collaborative input of all stakeholders—students, lecturers, and administrators—in negotiating new systems of teaching and learning with stronger lecturer buy-in in coming alongside students. This will require lecturers to upskill or risk becoming obsolete within the current and

future AI-infused HEI context. Moreover, open educational practices within the context of HEIs in the Global South have the capacity to promote themselves as democratised centres for transformative thinking as they utilise a new form of authentic teaching, learning, and assessment.

## Acknowledgements

This work is based on the research supported in part by the National Research Foundation of South Africa (Ref Number SRUG2204285127).

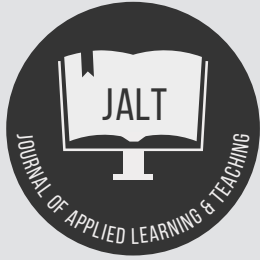
## References

- Ahmad, M., Subih, M., Fawaz, M., Alnuqaidan, H., Abuejheisheh, A., Naqshbandi, V., & Alhalaiqa, F. (2024). Awareness, benefits, threats, attitudes, and satisfaction with AI tools among Asian and African higher education staff and students. *Journal of Applied Learning and Teaching*, 7(1), 1-8. <https://doi.org/10.37074/jalt.2024.7.1.10>
- Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, 2(1), 41-49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Baker, T., & Smith, L. (2019). *Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges*. Nesta Foundation. [https://media.nesta.org.uk/documents/Future\\_of\\_AI\\_and\\_education\\_v5\\_WEB.pdf](https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf)
- Baškarada, S. (2014). Qualitative case studies guidelines. *The Qualitative Report*, 19(40), 1-25. <https://doi.org/10.46743/2160-3715/2014.1008>
- Bozkurt, A., & Sharma, R. C. (2023). Challenging the status quo and exploring the new boundaries in the age of algorithms: Reimagining the role of generative AI in distance education and online learning. *Asian Journal of Distance Education*, 18(1). <https://orcid.org/0000-0002-4520-642X>
- Brockman, G. (2023). *The inside story of ChatGPT's astonishing potential* [Video]. TED2023. [https://www.ted.com/talks/greg\\_brockman\\_the\\_inside\\_story\\_of\\_chatgpt\\_s\\_astonishing\\_potential/?language=en](https://www.ted.com/talks/greg_brockman_the_inside_story_of_chatgpt_s_astonishing_potential/?language=en)
- Chaka, C. (2024). Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. *Journal of Applied Learning and Teaching*, 7(1), 1-12. <https://doi.org/10.37074/jalt.2024.7.1.14>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education.



- International Journal of Educational Technology in Higher Education, 20(43), 1-18. <https://doi.org/10.1186/s41239-023-00411-8>
- Chugh, R., Turnbull, D., Cowling, M. A., Vanderburg, R., & Vanderburg, M. A. (2023). Implementing educational technology in higher education institutions: A review of technologies, stakeholder perceptions, frameworks, and metrics. *Education and Information Technologies*, 28, 1-27. <https://doi.org/10.1007/s10639-023-11846-x>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). *Chatting and cheating. Ensuring academic integrity in the era of ChatGPT*. EdArXiv. <https://doi.org/10.35542/osf.io/mrz8h>
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://psycnet.apa.org/doi/10.2307/249008>
- Davis, F. D. (1993). User acceptance of information technology: System characteristics, user perceptions and behavioral impacts. *International Journal of Man-Machine Studies*, 38(3), 475-487. <https://doi.org/10.1006/imms.1993.1022>
- Dede, C. J. (2009). *Comparing frameworks for 21st century skills* [online]. Harvard Graduate School of Education. [https://www.dpsgs.org/pdf/Comparing\\_Frameworks\\_for.pdf](https://www.dpsgs.org/pdf/Comparing_Frameworks_for.pdf)
- Fullan, M., Azorín, C., Harris, A., & Jones, M. (2023). Artificial intelligence and school leadership: Challenges, opportunities, and implications. *School Leadership & Management*, 1-8. <https://doi.org/10.1080/13632434.2023.2246856>
- Holmes, W., Iniesto, F., Anastopoulou, S., & Boticario, J. C. (2023). Stakeholder perspectives on the ethics of AI in distance-based higher education. *International Review of Research in Open and Distributed Learning*, 24(2), 96-117. <https://doi.org/10.19173/irrodl.v24i2.6089>
- Holmes, W., & Porayska-Pomsta, K. (2023). *The ethics of AI in education: Practices, challenges, and debates*. Routledge. <http://dx.doi.org/10.4324/9780429329067>
- Ifelebuegu, A. O. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning & Teaching*, 6(2), 385-392. <https://doi.org/10.37074/jalt.2023.6.2.2>
- Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIED) tools: The good, the bad, and the ugly. *Journal of Applied Learning & Teaching*, 6(2), 332-345. <https://doi.org/10.37074/jalt.2023.6.2.29>
- Jones, M., & Sheridan, L. (2015). Back translation: An emerging sophisticated cyber strategy to subvert advances in 'digital age' plagiarism detection and prevention. *Assessment & Evaluation in Higher Education*, 40(5), 712-724. <https://doi.org/10.1080/02602938.2014.950553>
- Ladha, N., Yadav, K., & Rathore, P. (2023). AI-generated content detectors: Boon or bane for scientific writing. *Indian Journal of Science and Technology*, 16(39), 3435-3439. <https://doi.org/10.17485/IJST/v16i39.1632>
- Lindgren, S. (2023). Introducing critical studies of artificial intelligence. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 1-19). Edward Elgar Publishing. <https://doi.org/10.4337/9781803928562.00005>
- Maphoto, K. B., Sevnarayan, K., Mohale, N. E., Suliman, Z., Ntsopi, T. J., & Mokoena, D. (2024). Advancing students' academic excellence in distance education: Exploring the potential of generative AI integration to improve academic writing skills. *Open Praxis*, 16(2), 142-159. <https://doi.org/10.55982/openpraxis.16.2.649>
- McQuillan, D. (2021). Finding your voice in academic writing. In D. McQuillan (Ed.), *Peer-led student handbook series* (Chapter 3). Technological University Dublin. <https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1003&context=researchporbk>
- Mollick, E. R., & Mollick, L. (2022). *New modes of learning enabled by AI chatbots: Three methods and assignments* [online]. SSRN Electronic Journal. <https://dx.doi.org/10.2139/ssrn.4300783>
- Moore M. G. (2013). *Handbook of distance education*. Routledge.
- Naidu, K., & Sevnarayan, K. (2023). ChatGPT: An ever-increasing encroachment of artificial intelligence in online assessment in distance education. *Online Journal of Communication and Media Technologies*, 13(1), e202336. <https://doi.org/10.30935/ojcm/13291>
- Ogata, H., Flanagan, B., Takami, K., Dai, Y., Nakamoto, R., & Takii, K. (2024). EXAIT: Educational eXplainable Artificial Intelligent Tools for personalized learning. *Research and Practice in Technology Enhanced Learning*, 19(19). <https://doi.org/10.58459/rptel.2024.19019>
- OpenAI. (2022, November 30). *Introducing ChatGPT*. OpenAI. <https://openai.com/blog/chatgpt>
- Oppenheimer, D. (2023, January 17). ChatGPT has arrived and nothing has changed. *Times Higher Education*. <https://www.timeshighereducation.com/campus/chatgpt-has-arrived-and-nothing-has-changed>
- Ormond, E. (2023, June 11). AI's risks and rewards: With great power comes great responsibility. *Daily Maverick*. <https://www.dailymaverick.co.za/opinionista/2023-06-11-ais-risks-and-rewards-with-great-power-comes-great-responsibility/>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2), 7. <https://doi.org/10.53761/1.20.02.07>
- Popenici, S. (2023). The critique of AI as a foundation for judicious use in higher education. *Journal of Applied Learning & Teaching*, 6(2), 378-384. <https://doi.org/10.37074/jalt.2023.6.2.4>
- Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023). A critical

- perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning & Teaching*, 6(2), 311-331. <https://doi.org/10.37074/jalt.2023.6.2.5>
- Qasim, A., El Refae, G. A., & Eletter, S. (2022). Embracing emerging technologies and artificial intelligence into the undergraduate accounting curriculum: Reflections from the UAE. *Journal of Emerging Technologies in Accounting*, 19(2), 155-169. <https://doi.org/10.2308/JETA-2020-090>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1), 41-56. <https://doi.org/10.37074/jalt.2023.6.1>
- Rudolph, J., Ismail, M. F., & Popenici, S. (2024). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. *Journal of University Teaching and Learning Practice*, 21(6). <https://doi.org/10.53761/54fs5e77>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sevnarayan, K., & Maphoto, K. B. (2024). Exploring the dark side of online distance learning: Cheating behaviours, contributing factors, and strategies to enhance the integrity of online assessment. *Journal of Academic Ethics*, 22, 51-70. <https://doi.org/10.1007/s10805-023-09501-8>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31-40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Thompson, P. (2012). Achieving a voice of authority in PhD theses. In K. Hyland, & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 119-133). Palgrave MacMillan. [https://doi.org/10.1057/9781137030825\\_8](https://doi.org/10.1057/9781137030825_8)
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 1-24. <https://doi.org/10.1186/s40561-023-00237-x>
- Venkatesh, V., & Davis, F. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Weinberger, D. (2007). *Everything is miscellaneous: The power of the new digital order*. Henry Holt and Company.
- Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: New demands for teachers' roles. *Heliyon*, e24289. <https://doi.org/10.1016/j.heliyon.2024.e24289>



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review

---

Chaka Chaka<sup>A</sup>

A

*Professor, University of South Africa, Pretoria, South Africa*

---

### Keywords

Academic and scientific writing;  
AI detection accuracy;  
AI detection reliability;  
AI detection tools;  
AI-generated text;  
higher education;  
human-written text;  
large language models (LLMs);  
review study.

---

### Abstract

The purpose of this study was to review 17 articles published between January 2023 and November 2023 that dealt with the performance of AI detectors in differentiating between AI-generated and human-written texts. Employing a slightly modified version of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol and an aggregated set of quality evaluation criteria adapted from A Measurement Tool to Assess systematic Reviews (AMSTAR) tool, the study was conducted from 1 October 2023 to 30 November 2023 and guided by six research questions. The study conducted its searches on eleven online databases, two Internet search engines, and one academic social networking site. The geolocation and authorship of the 17 reviewed articles were spread across twelve countries in both the Global North and the Global South. ChatGPT (in its two versions, GPT-3.5 and GPT-4) was the sole AI text generator used or was one of the AI text generators in instances where more than one AI text generator had been used. Crossplag was the top-performing AI detection tool, followed by Copyleaks. Duplichecker and Writer were the worst-performing AI detection tools in instances in which they had been used. One of the major aspects flagged by the main findings of the 17 reviewed articles is the inconsistency of the detection efficacy of all the tested AI detectors and all the tested anti-plagiarism detection tools. Both sets of detection tools were found to lack detection reliability. As a result, this study recommends utilising both contemporary AI detectors and traditional anti-plagiarism detection tools, together with human reviewers/raters, in an ongoing search for differentiating between AI-generated and human-written texts.

### Correspondence

chakachaka8@gmail.com<sup>A</sup>

---

### Article Info

Received 16 December 2023  
Received in revised form 18 January 2024  
Accepted 3 February 2024  
Available online 7 February 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.14>

## Introduction

Since the launch of ChatGPT on 30 November 2022, much research, both academic and non-academic papers, and numerous preprints have been published on the multiple uses for which generative artificial intelligence (AI) chatbots or AI-powered large language models (LLMs) can be put to educational settings. These types of chatbots are also known as AI text generators. The multifarious affordances of these AI text generators are now well documented. Some of these affordances include educational content generation (Ifelebuegu et al., 2023; Kasneci et al., 2023; Perkins et al., 2023; cf. Chaka, 2023a; Rudolph et al., 2023), enhancing online assessments and supporting collaborative assessments (Gamage et al., 2023; Ifelebuegu, 2023a; Kasneci et al., 2023), language learning and personalised learning (Chaka, 2023a, 2023b; Hew et al., 2023; Ifelebuegu et al., 2023; Jeon & Lee, 2023), student learning (Abbas et al., 2023; Hew et al., 2023; Sullivan et al., 2023); essay writing (Chaka, 2023a; Yeadon et al., 2023); student/teaching assistants (Jeon & Lee, 2023; Kasneci et al., 2023; Kuhail et al., 2023; Nah et al., 2023); and conducting and publishing research (Kooli, 2023; van Dis et al., 2023). Equally, the various challenges and risks AI chatbots pose in academia have now been profusely documented as well. Among these are academic dishonesty and plagiarism (Chaka, 2023a, 2023c; Cotton et al., 2023; Ifelebuegu, 2023; Ifelebuegu et al., 2023; Kasneci et al., 2023; Kleebayoon & Wiwanitkit, 2023; Kooli, 2023; Perkins et al., 2023; Rudolph et al., 2023; Sullivan et al., 2023) and bias and unfairness (Dwivedi et al., 2023; Kasneci et al., 2023; Nah et al., 2023; Ray, 2023). To this effect, some review studies have been conducted on the use of the new AI chatbots in education in general (see Baidoo-Anu & Ansah, 2023; Dergaa et al., 2023; Ifelebuegu et al., 2023; Perera & Lankathilaka, 2023; Pinzolit, 2023; Sullivan et al., 2023; Thurzo et al., 2023; Yang et al., 2023). For example, Baidoo-Anu and Ansah's (2023) study investigated, among other things, the potential benefits of ChatGPT in education as reported in peer-reviewed journal articles and/or in preprints published between November 2022 and March 2023. In addition, Dergaa et al.'s (2023) study explored the possible benefits and threats of ChatGPT and other natural language processing technologies in academic writing and research publications as reported in peer-reviewed journal articles indexed in Scopus's quartile 1.

Importantly, instances of AI tools that can detect AI-generated content and that can distinguish this type of content from the one written by humans have been investigated by a number of scholars. Scholars who have done so include, among others, Abani et al. (2023), Alexander et al. (2023), Anil et al. (2023), Chaka (2023c), Elkhayat et al. (2023), Gao et al. (2023), Perkins et al. (2023), and Uzun (2023). These scholars have done so in varying degrees and by focusing on different types of AI detection tools. The detection tools explored include single detection tools (Habibzadeh, 2023; Perkins et al., 2023; Subramaniam, 2023); two detection tools (Bisi et al., 2023; Desaire et al., 2023; Ibrahim, 2023); three detection tools (Cingillioglu, 2023; Elali & Rachid, 2023; Gao et al., 2023; Homolak, 2023; Ladha et al., 2023; Wee & Reimer, 2023); four detection tools (Abani et al., 2023; Alexander et al., 2023; Anil et al., 2023); and multiple detection tools (Chaka, 2023c; Odri & Yoon, 2023; Santra

& Majhi, 2023; Walters, 2023). But more scholarly papers published in this area are preprints, which, at the moment, tend to outnumber journal articles and book chapters. However, unlike the picture painted above, there are, if any, few review studies that have been published in this area (cf. Baidoo-Anu & Ansah, 2023; Dergaa et al., 2023; Ifelebuegu et al., 2023; Perera & Lankathilaka, 2023; Pinzolit, 2023; Sullivan et al., 2023; Thurzo et al., 2023; Yang et al., 2023). Rather, the bulk of scholarly papers published in this area are, again, preprints (see Aremu, 2023; Maddugo, 2023; Weber-Wulff et al., 2023) and, to some extent, conference proceedings (see Sarzaeim et al., 2023; Singh, 2023).

At the time of writing this paper, there was no published peer-reviewed review journal article on AI detection tools differentiating between AI-generated and human-written texts. Such review publications are essential for the purpose of framing a related work section to highlight and interrogate issues pertaining to specific AI detection tools that relevant review studies have explored. So, in the absence of such studies, the present paper will not have a related work section. This paper consists of the following sections: the purpose of the study, article characteristics and research questions, methods (search strategy, eligibility criteria and selection of peer-reviewed journal articles, quality evaluation, coding, and inter-rater reliability, data extraction and analysis), findings and discussion, and conclusion. All of these sections together constitute a review protocol (see Xiao & Watson, 2019).

## Purpose of the study, article characteristics, and research questions

In light of the points highlighted above, the purpose of this study was to review 17 articles published between January 2023 and November 2023 that focused on the performance of AI detection tools in differentiating between AI-generated and human-written texts. The focus of the study was on AI detection tools employed in the higher education (HE) sector during this period. This purpose was informed by the fact that the study wanted to establish which AI detection tools in the reviewed studies are reported to perform better in differentiating between AI-generated and human-written texts. Establishing which AI detection tools perform better and knowing whether their detection accuracy is reliable are some of the key factors confronting the HE sector since the release of ChatGPT and the proliferation of AI-powered chatbots that followed after its launch. The purpose of the study also had to do with the overall desire to contribute to review studies in this area of AI detection tools.

There were twelve article characteristics investigated in each review article. These were as illustrated in Table 1. To this end, the study had the following research questions (RQs):

- RQ 1: What types of articles have the current review study identified, and what discipline do they belong to?
- RQ 2: What is the purpose of each article?

- RQ 3: What are the AI-generated and human-written texts tested?
- RQ 4: What is the number and what are the names of the AI detection tools used, and what are the best- and worst-performing AI detection tools reported?
- RQ 5: What are the detection accuracy rate and the detection accuracy reliability reported?
- RQ 6: What are the main findings and the key conclusions of the 17 reviewed articles?

## Method

There are different typologies of review studies. For instance, Grant and Booth (2009) identify fourteen different types of review studies, of which rapid reviews, scoping reviews, literature reviews, systematic reviews, meta-analyses, and integrative synthesis reviews are but a few examples (cf. Xiao & Watson's 2019 sixteen types of review studies). These review types differ mainly in terms of their foci, aims, search strategies, appraisals, analyses, and syntheses (Grant & Booth, 2009). Due to space constraints, I will briefly describe a literature review and a synthesis review as they constitute the current study. The present study is a review that comprises literature and synthesis review components. In its literature review component, the study focused on currently published peer-reviewed journal articles on AI detection tools differentiating between AI-generated and human-written texts in more than one field of study. Its searches were comprehensive but constrained by a given timeline, and its quality assessment was proscribed by the scarcity of published peer-reviewed journal articles on its focus area.

Additionally, the study employed a thematic analysis and a narrative synthesis. In its integrative synthesis outlook, the study integrated and compared peer-reviewed journal articles currently published in its focus area and selected all the relevant articles that were retrievable from the online search platforms on which it conducted its search strategies. Similarly, its analysis and synthesis were thematic and narrative, respectively. Importantly, the aim of an integrative synthesis is to broaden how a given phenomenon is understood (see Grant & Booth, 2009; cf. Chaka, 2022, 2023d; Snyder, 2019; Xiao & Watson, 2019). When two different types of reviews have been fused, as is the case in this study, such a product is referred to as a hybrid review (see Xiao & Watson, 2019; also see Bacon, 2017). This type of hybrid review entails summarising and synthesising findings from reviewed studies (Bacon, 2017; Grant & Booth, 2009; Snyder, 2019).

Even though this study is a hybrid review study as specified above, for transparency purposes, it followed a slightly modified version of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol in its review process, as spelt out below. Four key features of the PRISMA reporting protocol are comprehensiveness, systematicity, transparency, and rigour in the review process (e.g., literature searches, screening and identifying eligible articles (publications), data extraction and analysis, and

summarising and synthesising findings) (see Chaka, 2022, 2023d; Ismail et al., 2023; Stracke et al., 2023; Yang et al., 2023).

## Search strategy

A literature search for potential peer-reviewed journal articles was carried out from 1 October 2023 to 30 November 2023. The search was conducted on Internet search engines, online databases, and one academic social networking site. These online search platforms were as follows: two Internet search engines (Google search and Microsoft Bing search), eleven online databases (Google Scholar, Semantic Scholar, Taylor & Francis Online, Wiley Online Library, ScienceDirect, Scopus, SpringerLink, IEEE Xplore Digital Library, ERIC, JSTOR, and BASE), and ResearchGate. Altogether, these constituted fourteen online platforms (see Figure 1; cf. Chaka, 2022, 2023d; Ismail et al., 2023; Stracke et al., 2023). All of these online platforms were easily accessible, while the others, such as EBSCO and Web of Science, had paywalls.

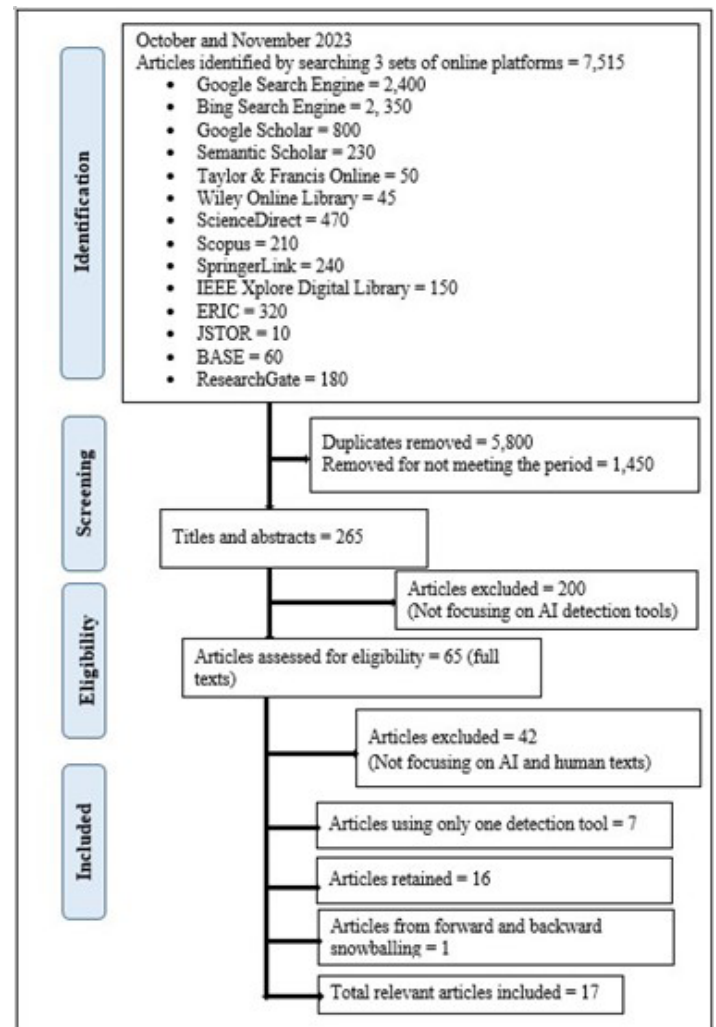


Figure 1: PRISMA flowchart for screening articles.

Search strings included keywords, phrases, and short clauses related to the focus area of the study: AI detection tools used in differentiating between AI-generated and human-written texts. Even though the application context of these AI detection tools was higher education, the search strings were left open-ended in order to source wide-ranging AI

detection tools. This was after the researcher had realised the scarcity of peer-reviewed journal articles focusing on this area at the time of conducting the present study. The search strings consisted of Boolean operators (AND or OR) (see Chaka, 2023d) and truncation symbols such as \*, \, or -, depending on the search platform. Moreover, the permutations of these search strings were used iteratively. Below are examples of the search strings that were employed:

- Published papers on AI-generated content detection tools in 2023 (Google, Bing, Google Scholar, and ResearchGate)
- Tools for detecting artificial intelligence-generated content (Taylor & Francis Online, Wiley Online Library, ScienceDirect, Scopus, and SpringerLink)
- Detecting AND AI texts OR human texts – 2023 (Google, Bing, Google Scholar, and ResearchGate)
- Differentiating between AI-generated and human-written text using AI detection tools (Semantic Scholar)
- Tools to detect AI-written and human-written text (Semantic Scholar)
- Differentiating between AI-generated and human-written text using AI detection tools (Wiley Online Library)
- Best AI tools to detect AI plagiarism; Plagia\* detec\*; Detect\* artificial intellig\* gener\* cont\*; Detect\* artificial intellig\* gener\* text; Detect\* tools artificial intellig\* gener\* text; Best artificial intellig\* tools for detect\* artificial intellig\* gener\* text; Artificial intellig\* tools for detect\* artificial intellig\* gener\* text; Artificial intellig\* detect\* tools (IEEE Xplore Digital Library).

### Eligibility criteria and selection of peer-reviewed journal articles

The eligibility criteria used to judge the suitability and relevance of the peer-reviewed journal articles for this study were based on the classical inclusion/exclusion format (see Chaka, 2022, 2023d; Ismail et al., 2023; Stracke et al., 2023). For example, the time-period inclusion criterion was peer-reviewed journal articles published between January 2023 and November 2023 (see Table 1). Eligible journal articles were determined through a search and screening process that was conducted on the fourteen aforementioned online search platforms during the specified coverage time frame. During this process, 7,515 articles were returned by the fourteen online search platforms (see Figure 1). Of these articles, 5,800 were duplicates and were removed while 1,450 did not meet the designated coverage time frame and were, also, accordingly, eliminated. The remaining articles (n = 265) were screened by reviewing their titles and abstracts. After this screening process, 200 articles were excluded as they did not focus on AI detection tools. A full-text review of the remaining 65 articles was conducted, after which 41

articles were eliminated due to their lack of focus on AI and human texts. Of the remaining 23 articles, 7 articles were excluded as they each used only one AI detection tool for distinguishing between AI-generated and human-written texts. This led to 16 articles being retained (see Figure 1).

Table 1: Inclusion/exclusion criteria.

Criteria	Inclusion	Exclusion
Time period	Peer-reviewed journal articles published between January 2023 and November 2023	Peer-reviewed journal articles not published between January 2023 and November 2023
Types of articles	Articles published in peer-reviewed journals	Articles not published in peer-reviewed journals
Content and focus of articles	Peer-reviewed journal articles focusing on AI detection tools that can differentiate between AI-generated and human-written texts	Peer-reviewed journal articles not focusing on AI detection tools that can differentiate between AI-generated and human-written texts
Number of AI detection tools	More than one AI detection tool, one of which is meant to detect the currently available LLM chatbots*	Only one AI detection tool meant to detect the currently available LLM chatbots
Language of publication	Peer-reviewed journal articles published in English	Peer-reviewed journal articles not published in English

\* LLM chatbots released after ChatGPT or released to compete with it.

From the 16 retained articles, forward snowballing, and backward snowballing – also known as descendent and ancestry searches – were conducted to further identify suitable and eligible articles (see Chaka, 2022, 2023d; Wohlin et al., 2022). Forward snowballing entails searching and locating publications that cite the publications established during the search process; backward snowballing involves searching and locating publications listed in the reference lists of publications discovered during initial literature searches (see Chaka, 2022; Wohlin et al., 2022). The resultant dual snowballing search yielded one more relevant and eligible article. Overall, then, the total number of suitable and eligible articles for the present study was 17 (see Figure 1). The reviewing of the 17 articles was done manually.

### Quality evaluation, coding, and inter-rater reliability

Evaluating and ensuring methodological quality is essential for review studies. This is so even when there is a scarcity of review studies in any given area of focus. There are quality assessment criteria recommended by scholars such as Kitchenham et al. (2009) and Shea et al. (2009). The present review study formulated and utilised an aggregated set of quality evaluation criteria adapted from A Measurement Tool to Assess systematic Reviews (AMSTAR) tool (Shea et al., 2009; Shea et al., 2017; also see Chaka, 2022; Li et al., 2022) and from the quality evaluation guidelines designed by Kitchenham et al. (2009) and Kitchenham and Brereton (2013). Based on these sixteen quality evaluation criteria, a checklist form was formulated (see Table 2). However, since this is not a systematic literature review, and as there was a dearth of peer-reviewed articles published in the focus area of this study, as mentioned earlier, the quality evaluation criteria used here are customised for this study, even though they have some universal applicability for review studies on AI detection tools. The application of the quality evaluation checklist was not rigid but flexible.

Concerning the reviewed articles, two raters (including the author of this article) independently evaluated each article using the checklist illustrated in Table 2. A “yes or “no” rating was allocated to each of the sixteen criteria for each

Table 2: Quality evaluation questions.

Article quality evaluation questions	
1.	Are the aims/purposes of the article clearly stated?
2.	Is the field of study (subject area) of the article mentioned?
3.	Are the genres of the AI and human texts tested provided?
4.	Is there a specific method/methodology specified?
5.	Is there more than one AI detection tool used?
6.	Are the names of the AI tools used to generate texts specified and are the protocol, procedure, prompts, re-prompting, revised prompts, etc., for generating AI texts mentioned?
7.	Is the nature of the humans used to generate texts specified mentioned?
8.	Is the justification for the choice of the AI tools to generate texts given?
9.	Are the names of the AI tools used to detect AI-generated and human-written texts specified?
10.	Is the justification for the choice of the AI tools used to detect AI-generated and human-written texts given?
11.	Are the AI-generated and human-written texts for detection by AI tools both sufficient and credible in terms of how they have been described?
12.	Is there transparency in the way the AI-generated and human-written texts were collected?
13.	Is there transparency in reporting the way in which the AI-generated and human-written texts were subjected to the AI detection tools used?
14.	Are the findings grounded on the data and credible?
15.	What contribution do the findings make to the existing knowledge or understanding?
16.	Are the conclusions based on the findings?

article, with a “yes” rating allotted the number 1 (one) and a “no” rating assigned the number 0 (zero). The two raters’ rating agreement scores were calculated following Cohen’s kappa coefficient ( $\kappa$ ) (see Cohen, 1960). Rating discrepancies between the two raters were resolved by discussing them and by reaching a consensus (Landis & Koch, 1977; Pérez et al., 2020). The inter-rater agreement was calculated using Landis and Koch’s (1977) scoring and its related interpretation. The inter-rater agreement represents the extent of autonomy raters exhibit in scoring items by attempting to reach the same conclusion. Using Landis and Koch’s (1977)  $\kappa$  values of  $<0$  = poor,  $0.00-0.20$  = slight,  $0.21-0.40$  = fair,  $0.41-0.60$  = moderate,  $0.61-0.80$  = substantial, and  $0.81-1.00$  = almost perfect, which are modifications of Cohen’s (1960) original labels, the inter-rater agreement between the two raters was 0.82. As this joint agreement score falls within the 0.81-1.00 almost-perfect score range, it was deemed acceptable (also see Chaka, 2022, 2023d; McHugh, 2012).

### Data extraction and analysis

Based on the quality evaluation criteria, the coding procedure, and the inter-rater reliability described above, datasets were extracted from the peer-reviewed articles included in this study. These datasets were in the form of the twelve journal characteristics illustrated in Table 3. This table also served as an analytic scheme for thematic analysis that was conducted on the extracted datasets. Categories and themes that responded to the research questions for this study were developed from this analysis (see Chaka, 2022, 2023d).

Table 3: Twelve key journal characteristics investigated in each review study.

Author(s) and year and month of publication	Number and names of AI detection tools
Country of origin	Best and worst performing AI detection tool(s)
Article type	Detection accuracy rate
Discipline/subject area	Detection accuracy reliability
Purpose	Main findings
AI and human texts tested	Key conclusions

### Findings and discussion

The findings presented in this part of the paper are based on the datasets extracted from the 17 selected journal articles. They are presented according to the twelve journal characteristics and in line with the research questions (RQs) mentioned earlier. These findings are integrated with their discussion.

### Authors, countries of origin, article types, disciplines, and purposes

The 17 reviewed articles were produced by authors from twelve countries: India, the USA, Germany, Greece, France, South Africa, Australia, Hong Kong, Qatar, Croatia, Kuwait, and Malaysia. Three articles were written by authors from two countries: India and the USA. Two articles were produced by authors from France. The remaining articles ( $n = 9$ ) were written by authors from nine different countries (see Table 4). At a geolocal vantage point, there is an infinitesimal difference between the number of articles contributed by countries deemed to represent the Global North and those by countries viewed to represent the former block of countries have over the latter block in this review study. This geolocal and authorship distribution, which is often viewed as a proxy for the geopolitics and economy of knowledge production (see Chaka, 2023e; Müller, 2021; R’boul, 2022; also see Domínguez et al., 2023), seems not to resonate with the views and findings of Chaka (2023e), Müller (2021), and R’boul (2022), at least in the context of this study. While this does not invalidate or deny the views and findings of these scholars’ studies, as their contexts and dynamics differ, the current study articulates one of the observations that emanates from it. Without denying the existence of the geopolitics of knowledge and of the geospatial entanglements of knowledge, this observation is instructive, though, since the study did not use *geopolitics and economy of knowledge production nor names of countries in its search strings*.

Table 4: Article numbers, types, authors, countries, texts tested, AI tools used, the best and worst performing AI tools.

Word performing tool	Turbin	Open AI Classifier	DupliChecker	NOW	GLTR	Open AI Classifier	Open AI Classifier	NOW	Writer	iThenticate	NOW	Crosscap	Content at Scale	GPTZero & DupliChecker	Original iThenticate & DupliChecker	Spelling & Grammar	GPT-3 Detector	
Best performing tool	Open AI Classifier	Crosscap	Grammarly	NOW	Copyleaks	GPTZero	ZensGPT	NOW	Crosscap	GPT-3 Output Detector	NOW	GPT-3 Output Detector	Writer	Originality	Content at Scale & Spelling	Turnitin	Content at Scale	
Number AI detection tools used	4	4	4	2	5	3	2	3	5	3	3	2	3	11	8	16	3	
Texts tested	Two sets of 3 research papers (A&H)	Four academic writing essays (A&H)	200 essays (A)	425 original & MT articles	Three sets of AI English texts & MT texts (n = 21)	150 essays (75 + A; 75 = human)	300 paragraphs (100 = human, 200 = AI)	A fabricated abstract	15 paragraphs (A)	100 abstracts, 200 paragraphs (100 = human, 100 = AI)	240 essays (A&H)	Four research articles (H&A)	8 texts (7 = human, 1 = AI)	80 academic writing samples (A)	42 = human, 84 = AI	138 essays (A)	18 essays (H)	
Country	Germany	Greece	India	France	South Africa	Australia	Hong Kong	USA	Qatar	USA	Croatia	Horobaki	Iran	India	France	India	USA	Malaysia
Author(s)	Allen et al. (2023)	Alexander et al. (2023)	Anil et al. (2023)	Bisai et al. (2023)	Chaka (2023)	Ongilagus (2023)	Desaire (2023)	Esai & Rachid/Elkhatir et al. (2023)	Gao et al. (2023)	Gao et al. (2023)	Horobaki (2023)	Iranian (2023)	Laitha et al. (2023)	Oth & Yoon (2023)	Sentra & Nagesh (2023)	Waters (2023)	Woo & Raner (2023)	
Article number & type	Art. 1 (OP)	Art. 2 (RP)	Art. 3 (RP)	Art. 4 (RP)	Art. 5 (RP)	Art. 6 (RP)	Art. 7 (Report)	Art. 8 (RP)	Art. 9 (OA)	Art. 10 (BC)	Art. 11 (Comment)	Art. 12 (RP)	Art. 13 (RP)	Art. 14 (RP)	Art. 15 (RP)	Art. 16 (RP)	Art. 17 (RP)	

Notes: NOW = no outright winner; A&H = artificial intelligence-generated & human-written; MT = machine translated translation; OP = opinion paper; RP = research paper; OA = original article; BC = brief communication; comment = commentary; VP = viewpoint

NB: The early online versions of Bisai et al. (2023) and Oth and Yoon (2023) are the ones that were reviewed. Their latest versions were published in December after the study had been conducted. Except for their bibliographic details, the contents of these articles are the same in both versions.

The articles reviewed in this study were of different types: research papers ( $n = 11$ ), opinion papers ( $n = 2$ ), commentary ( $n = 1$ ), report ( $n = 1$ ), brief communication ( $n = 1$ ), and viewpoint ( $n = 1$ ) (see Table 4). All of these articles were published or available online between 10 March 2023 and 15 November 2023, with three articles published in April and October, respectively (see Figure 2). Research articles, or original papers, predominated the other article types. This is an unexpected but not a surprising development since most AI-related scholarly papers, including scholarly papers on AI detection tools that can differentiate between AI-generated and human-written texts, were an instantaneous response to ChatGPT after its viral launch on 30 November 2022. This particular development tends to resemble, albeit for different

reasons and for dissimilar dynamics, the exponential growth in the number of scholarly papers and preprints that were published immediately after the outbreak of the COVID-19 pandemic. During this period, too, many commentaries, reports, and viewpoints were instantly published (see Chaka, 2020).

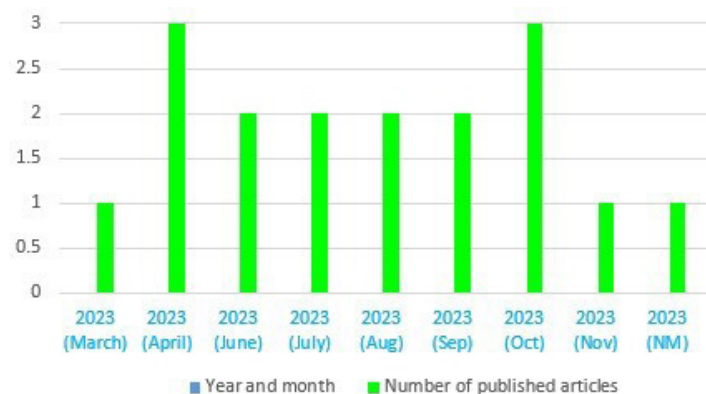


Figure 2: Number of published articles and years and months of publication.

The academic disciplines covered by the 17 reviewed articles are many and varied. Medical and biomedical sciences, together with hard sciences (e.g., chemistry), dominated, followed by English language studies (see Table 5). This observation should be seen against the backdrop of the disciplines in which ChatGPT and the other generative AI models seem to pose the biggest threat in terms of academic integrity. For medical and biomedical sciences and hard sciences, it is the integrity of scientific writing that generative AI models like ChatGPT threaten. Similarly, for English language studies, which have academic essay writing and composite studies as some of its flagship assessment methods, the emergence of generative AI models is not a transient fad: it is a big issue that goes to the heart of its existence. So, for all of these disciplines, testing or evaluating which detection tools can discriminate between AI-generated and human-written texts with the highest accuracy and the maximum reliability is a matter of life and death (see Kenwright, 2024; Uzun, 2023; cf. Lim et al., 2023).

Table 5: Articles, disciplines and purposes.

Articles	Disciplines	Purposes
Art. 1	Veterinary neurology	Exploring the possible advantages and limitations of ChatGPT in scientific writing related to veterinary neurology
Art. 2	English as a Second Language (ESL)	Sharing information on the challenges confronted by English as a Second Language (ESL) lecturers in identifying AI-(ChatGPT)-generated texts.
Art. 3	Ten Miscellaneous fields of study: animal use, cosmetics and pharmaceutical industry, cosmology, engineering, environment, evolution of sports, finance, gender roles, medical technology, and chronic diseases.	Comparing overall similarity index (OSI) of four plagiarism detection tools and evaluating the factors affecting their effectiveness in detecting plagiarism.
Art. 4	Orthopaedics and traumatology	Screening for AI-generated content in articles published in Orthopaedics & Traumatology: Surgery & Research (OTSR) before and after the release of ChatGPT.
Art. 5	Applied English language studies (AELS)	Testing the accuracy of five AI content tools, GPTZero, OpenAI Text Classifier, Writer.com's AI Content Detector, CopyLeaks AI Content Detector, and Giant Language model Test Room, in detecting AI-generated content.
Art. 6	Data-driven predictive policing	Discussing the method and strategies to maintain academic integrity in educational settings.
Art. 7	Physical science (Chemistry)	Accurately detecting AI text when ChatGPT is told to write like a chemist.
Art. 8	Orthopaedics and rheumatology	Determining how AI-generated chatbots can be used to fabricate research in the medical community; and testing the accuracy of free, online AI detectors.
Art. 9	Chemical engineering/ Chemistry	Investigating the capabilities of various AI content detection tools in distinguishing between human and AI-authored content.
Art. 10	Biomedical sciences	Determining if ChatGPT can write convincing medical research abstracts.
Art. 11	Pharmacology	Examining the use of ChatGPT in academic writing and to evaluate the dependability of existing AI detectors.
Art. 12	English as a Second Language (ESL) writing	Examining the potential of two AI-based classifiers to detect AI-assisted plagiarism in ESL composition classes.
Art. 13	Scientific writing	Assessing the AI detection sites on a text generated wholly by the AI; testing the methods provided for evading AI detectors.
Art. 14	Orthopaedics and traumatology	Assisting publishers to identify AI-generated text in scientific research, academic work, and assignments as a means of regulating and promoting the ethical usage of AI in academia.
Art. 15	Library and information science	Examining the capability and shortcomings of typical plagiarism detectors to identify machine-generated scholarly text.
Art. 16	Social sciences, the natural sciences, and the humanities	Evaluating the accuracy of 16 AI text detectors in distinguishing between AI-generated and human-generated writing.
Art. 17	Bioscience	Examining the detection limits of AI detection tools to identify human-written essays, ChatGPT-generated essays, and partially AI-aided essays by checking AI procedurally generated essays in English and three other languages (Malay, Mandarin, Japanese, and their AI-translated versions).

Note: For full versions of some of these purposes, see Appendix A.

Moreover, the 17 articles had their specific purposes. While these purposes appear to be many and divergent, the convergence point is examining, evaluating, assessing, or testing the capabilities, potential, or accuracy and shortcomings or limits of AI detection tools in identifying AI-generated texts or in distinguishing between AI-generated and human-written texts in varying degrees (see Table 5). Two articles' (Art. 6 and Art. 10) purposes are generic. However, the purpose of Art. 6 is to preserve academic integrity by utilising AI detection tools in higher education. All of these purposes are about the detection of and the differentiation between AI-generated and human-written content as mediated largely by the AI detection tools utilised by the respective articles. Elsewhere, one of the purposes of Maddugoda's (2023) paper, which has some resonance with the convergence points of the purposes of the 17 articles, was to assess the efficacy of traditional anti-plagiarism software tools against some of the current AI detectors in identifying AI-generated content.

### The AI-generated and human-written texts tested

Twelve of the 17 articles utilised both AI-generated and human-written texts, with ChatGPT as a common text generator in all of them. Some of them had varying versions of ChatGPT-generated texts, such as original, fabricated, slightly modified, paraphrased or translated versions. Five articles employed only AI-generated texts. Of these, three articles employed ChatGPT as their preferred AI text generator (see Art. 3, Art. 8, and Art. 15). One article (Art. 9) used both ChatGPT-3.5 and ChatGPT-4, while another (Art. 5) utilised ChatGPT, YouChat, and Chatsonic. In their paper, Wu et al. (2023) provide large language model- (LLM) generated and human-written datasets that can be used as test datasets for detecting LLM-generated and human-written texts. Among these datasets are ChatGPT- or AI-generated datasets that can serve as the basis for comparing ChatGPT-generated text with human-written text. Likewise, Weber-Wulff et al.'s (2023) paper compared AI detection tools that could reliably distinguish between ChatGPT-generated texts and human-written texts. In these two studies, as is the case with the current study, ChatGPT-generated text serves as one of the pieces of AI-generated text.

### Number and names of the AI detection tools used

The number of AI detection tools employed by the 17 reviewed articles ranged from two to sixteen. Six articles (Art. 6, Art. 8, Art. 11, Art. 13, Art. 17, Art. 20) each used three AI tools, while three articles utilised two AI detection tools and four AI detection tools, respectively. Only two articles employed five AI tools. The remaining articles tested eight, eleven, and sixteen AI detection tools apiece. In this case, articles that employed three AI tools predominated. A paper that compared three AI tools is Singh's (2023), whereas Weber-Wulff et al.'s (2023) paper tested fourteen AI detectors.



## **Best- and worst-performing AI detection tools reported, and the detection accuracy rate and the detection accuracy reliability reported**

Concerning the best-performing AI detection tools, OpenAI Text Classifier, Crossplag, Grammarly, Copyleaks, for Art. 1, Art. 2, Art. 3, and Art. 5, respectively, had a better detection accuracy than their counterparts. The same is the case for Originality and Crossplag, Content at Scale and Sapling, Copyleaks and Turnitin, and Content at Scale in Art. 14, Art. 15, Art. 16, and Art. 17, correspondingly. With regard to Art. 6, GPTZero's detection accuracy was fractionally better than that of Copyleaks, while Crossplag had a marginal advantage over the other four detection tools in terms of detection accuracy in Art. 9. What is noteworthy is that in the case where eleven AI detection tools were tested, Originality and Crossplag did fairly better than the other nine tools. And, where sixteen AI detection tools were evaluated, Copyleaks and Turnitin had a higher detection accuracy than the other fourteen detectors. At a simple numerical level, Crossplag can be regarded as the best-performing AI detection tool as it topped or as it was one of the top-performing tools in at least three of the 17 reviewed articles (see Art. 2, Art. 9, and Art. 14). It is followed by Copyleaks that topped and co-topped in Art. 5 and Art. 16, respectively.

Concerning the other reviewed articles, the AI detection tools they tested either had a low detection accuracy (see Art. 4, Art. 5, Art. 7, Art. 8, and Art. 10), or displayed inconsistencies in their detection accuracy (see Art. 9, Art. 11, Art. 12, and Art. 13). Two AI detection tools that had tended to perform badly in the two instances (articles) in which had been used are Duplichecker (Art. 3 and Art. 15) and Writer (Art. 5 and Art. 9).

However, a word of caution is needed here. Notwithstanding the fact that some of the aforesaid AI detection tools did better than their counterparts as indicated above, they, nevertheless, fared badly in the other instances in which they were tested in some of the reviewed articles. For instance, the following AI detectors did badly in the reviewed articles indicated in parentheses: OpenAI Text Classifier (Art. 2, Art. 6 & Art. 7); Crossplag (Art. 12); Content at Scale (Art. 13); Sapling (Art. 16); and GPTZero (Art. 14). This points to some inconsistencies in these detection tools' accuracy when it comes to differentiating between AI-generated and human-created texts. Elkhatat et al. (2023) highlight this inconsistency bluntly when referring to the five AI detection tools they tested (see Art. 9) by opining that their performance was not completely reliable. This is because the AI detection tools they tested were inconsistent: they correctly identified some of the content of control responses (human-created texts) as having not been AI-generated while simultaneously displaying false positives and undecided classifications for the other portions of the same content. In fact, Wu et al. (2023) contend that none of the current state-of-the-art AI detection tools is infallible. In particular, the detection efficacy of AI detectors gets reduced by adversarial attacks, which are techniques or attempts to deliberately modify, fabricate, or manipulate text that goes beyond simple prompts (see Sayeed, 2023). For example, AI detectors are eluded by tampering with punctuation marks (e.g., removing a comma) in a text, and by applying synonym

substitution, paraphrasing/rewording, and translating a text (Wu et al., 2023; also see Krishna et al., 2023). In addition, they can be tricked by instances of single spacing (Cai & Cui, 2023; also see Chaka, 2023c). Moreover, most of the current AI detectors do not perform well in multilingual texts due to their monolingual AI detection algorithms (see Chaka, 2023c; Wu et al., 2023).

So, if reliability is construed to refer to any AI detection tool's capability to consistently detect AI-generated text with 100% precision (with no false positives) and human-written text with 100% precision (with no false negatives) across all contexts of writing, then, all reviewed AI detectors in this study cannot be regarded as reliable as none of them met this reliability requirement. Most crucially, because of their varying degrees of inconsistency in their detection efficacy as pinpointed in the preceding paragraph, all of them were highly unreliable. This aspect, again, brings into sharp focus Wu et al.'s (2023) view that the currently available AI detectors are fallible. This view resonates with Chaka's (2023c) contention that most of the current AI detection tools are not yet fully ready to convincingly and accurately detect AI-generated content from machine-generated texts in different domains. Actually, Sayeed (2023) goes on to assert that detecting AI-generated text in a reliable way is increasingly becoming mathematically impossible for the current AI detection tools. Given the findings of the present review study, I am strongly persuaded to concur with this contention. While on this issue of AI detection unreliability and inaccuracies, it is worth mentioning that OpenAI, the company behind ChatGPT-3.5 and GPT-4, is reported to have quietly discontinued its own AI detection tool, OpenAI Text Classifier, due to its detection unreliability and inaccuracies. It is reportedly mulling over bringing a better version of its AI detection tool (see Dreibelbis, 2023) back to business.

## **Main findings and key conclusions**

Some of the main findings of the reviewed articles touted the opportunities– potential solutions – offered by LLMs like ChatGPT, while flagging the challenges or threats posed by LLMs, especially in the area of academic and scientific writing. The opportunities relate to how such LLMs can benefit non-native English speakers in enhancing their academic and scientific writing (see Art. 2 and Art. 17). However, the catch is the plagiarism and the scientific dishonesty that LLMs encourage for academic and scientific writing (see Art. 2, Art. 13, Art. 15, and Art. 17). This set of main findings reflects how LLMs are double-edged or Janus-faced AI tools, at least for now. This is not a new observation, though. Well before the advent of ChatGPT, a paper by Sumakul et al. (2022) explored whether AI was a friend or a foe in English in foreign language (EFL) classrooms. After the release of ChatGPT, many papers have been published highlighting the benefits and challenges of ChatGPT in higher education. One such paper is Rasul et al.'s (2023). The other set of main findings concerns the inconsistencies of the AI detection tools tested in accurately and reliably distinguishing between AI-generated and human-written text. More than half of the reviewed articles reported on the inconsistencies of the AI detection tools they tested in their main findings (see Art. 1, Art. 3, Art. 4, Art. 5, Art. 9, Art. 10, Art. 11, Art. 13,

The detection inconsistencies of the AI detectors used in the reviewed articles have been dealt with and contextualised in the preceding section. Suffice it to say that one article (Art. 15) had as part of its main findings the fact that traditional anti-plagiarism tools (e.g., Turnitin, Grammarly, iThenticate) lack the ability to detect AI-generated text due to the differences in syntax and structure between machine-generated and human-written text. Dalalah and Dalalah (2023) take this shortcoming a step further by pointing out that discriminating between AI-generated text and simply copied text is rather difficult as AI detection algorithms are merely configured to detect whether a given text is machine-generated or not. A rider needs to be added to this point. AI detectors can only determine whether a text is AI-generated or not: they cannot establish the originality of a text even if it is copied. Doing so is the province of anti-plagiarism detection tools such as Turnitin, Grammarly, and iThenticate. The irony of anti-plagiarism detection tools, however, is that they do not necessarily detect plagiarism, but, rather, similarity indices. Added to this is the finding of Art. 15, which seems to loom large over them. Differentiating between and detecting plagiarised text and copied text, in addition to differentiating between and detecting AI-generated text and human-written text, is likely to become an even murkier minefield for AI detection tools as Microsoft's generative AI assistant, Microsoft 365 Copilot, is ready to be integrated into Microsoft 365 apps such as Word, Outlook, Teams, Excel, and PowerPoint. A similar generative AI assistant is likely to be integrated into the Google suite comprising Gmail, Docs, Slides, and Forms by Google (see Finnegan, 2023). While this generative AI integration might be beneficial for text predicting and for automating writing (e.g., drafting emails and creating slideshows), its downside is its potential to make up facts (hallucinate) and to spew inaccurate and false information (see Finnegan, 2023). All of this, then, adds another layer of AI-generated writing that AI detection tools will need to contend with in addition to simply differentiating between AI-generated and human-written texts.

Pertaining to the key conclusions, one set flagged the fact that the detection capability of most AI detection tools is largely confined to English (see Art. 4, Art. 5, Art. 9, Art. 12, Art. 17). The inability of some of the current AI detectors to function in texts written in other languages than English (including major European languages) is raised by, among others, Chaka (2023c) and Wu et al. (2023). For instance, Wu et al. (2023) argue that the main current AI detectors are designed to detect pieces of LLM-generated text meant for monolingual, and not multilingual, applications (also see Wang et al., 2023). Another key conclusion reported in this study is the need to use more than one AI detection tool, while another key conclusion is that AI detection tools need to be complemented by human reviewers. To add to these two points, in the unfolding environment of rapidly increasing AI text-generation tools and their attendant refinement, I think there is a need to employ a set of AI detection tools comprising traditional anti-plagiarism detection tools and AI detectors, on the one hand, and to enlist human reviewers/raters, on the other hand, for purposes of distinguishing between AI-generated text and

human-written text.

Finally, the other key conclusions are about the need for more development and refinement of AI content detection tools, the necessity to provide digital literacy training for teachers/human raters, and the need for journals to review their existing evaluation policies and practices in the light of AI. All of this calls for doing things differently across all domains, especially in academia, in the era of LLMs like ChatGPT.

## Conclusion

This study set out to review 17 articles published between January 2023 and November 2023 that dealt with the performance of AI detectors in differentiating between AI-generated and human-written texts. It was guided by six research questions (RQs). Authors from twelve countries wrote the reviewed articles. Viewed within its context, the geolocal and authorship dispersion of these articles tend not to dovetail with the geopolitics and economy of knowledge production as advanced by scholars such as Chaka (2023e), Müller (2021), and R'boul (2022). While the reviewed articles were of diverse types, the predominant article types were research papers, a finding that suggests that within less than a year after the release of ChatGPT, there were already studies conducted on AI detection tools that could distinguish between AI-generated and human-written texts. Among the academic disciplines explored, medical and biomedical sciences, together with hard sciences, dominated. They were followed by English language studies.

Even though the purposes of the 17 articles were many and varied, they converged in terms of examining, evaluating, assessing, or testing the capabilities, potential, or accuracy and shortcomings or limits of AI detection tools in identifying AI-generated texts or in differentiating between AI-generated and human-written texts in different contexts. The types of texts evaluated by these articles were AI-generated and human-written texts or AI-generated texts. In these two sets of texts (the former and latter text sets), ChatGPT (in its two versions, GPT-3.5 and GPT-4) was the sole AI text generator used or was one of the AI text generators in instances where more than one AI text generator had been used. The lowest number of AI detection tools was two, whereas the highest number of AI detection tools was sixteen. The names of the AI detectors used are displayed in Table 4.

In relation to the best-performing AI detection tools, Crossplag topped the other AI detectors in the three articles (Art. 2, Art. 9, and Art. 14) in which it had been tested. Copyleaks did so in two articles (Art. 5 and Art. 16). This finding should be seen in its context – the context of the 17 reviewed articles in this study as different AI detection tools tend to be prone to inconsistencies in the different contexts in which they are tested. Regarding the worst-performing AI detection tools, both Duplichecker and Writer fared badly in the articles in which they had been tested. However, the same caveat provided for the best-performing AI detectors above applies to them as well.

Lastly, one major aspect flagged by the main findings of the 17 reviewed articles is the inconsistency of the detection efficacy of all the tested AI detectors and all the tested anti-plagiarism detection tools. To this end, both sets of AI detection tools lacked detection reliability. Owing to this AI detection deficiency and the AI detection unreliability, the current study recommends employing both contemporary AI detectors and traditional anti-plagiarism detection tools, together with human reviewers/raters, in the pursuit of differentiating between AI-generated and human-written texts.

## References

Abani, S., Volk, H. A., De Decker, S., Fenn, J., Rusbridge, C., Charalambous, M., ... Nessler J. N. (2023). ChatGPT and scientific papers in veterinary neurology; Is the genie out of the bottle? *Frontiers in Veterinary Science*, *10*(1272755), 1-7. <https://doi.org/10.3389/fvets.2023.1272755>

Abbas, N., Ali, I., Manzoor, R., Hussain, T., & Hussaini, M. H. A. (2023). Role of artificial intelligence tools in enhancing students' educational performance at higher levels. *Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN)*, *3*(5), 36-49. <https://doi.org/10.55529/jaimlenn.35.36.49>

Alexander, K., Savvidou, C., & Alexander, C. (2023). Who wrote this essay? Detecting AI-generated writing in second language education in higher education. *Teaching English with Technology*, *23*(20), 25-43. <https://doi.org/10.56297/BUKA4060/XHLD5365>

Anil, A., Saravanan, A., Singh, S., Shamim, M. A., Tiwari, K., Lal, H., ...Sah, R. (2023). Are paid tools worth the cost? A prospective cross-over study to find the right tool for plagiarism detection. *Heliyon*, *9*(9), e19194, 1-11. <https://doi.org/10.1016/j.heliyon.2023.e19194>

Aremu, T. (2023). *Unlocking Pandora's box: Unveiling the elusive realm of AI text detection*. [https://papers.ssrn.com/sol3/Delivery.cfm/SSRN\\_ID4470719\\_code5947956.pdf?abstractid=4470719&mirid=1&type=2](https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID4470719_code5947956.pdf?abstractid=4470719&mirid=1&type=2)

Bacon, C. K. (2017). Multilanguage, multipurpose: A literature review, synthesis, and framework for critical literacies in English language teaching. *Journal of Literacy Research*, *49*(3), 424-453. <https://doi.org/10.1177/1086296X17718324>

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52-62.

Bisi, T., Risser, A., Clavert, P., Migaud, H., & Dartus, J. (2023). What is the rate of text generated by artificial intelligence over a year of publication in orthopedics and traumatology: Surgery and research? Analysis of 425 articles before versus after the launch of ChatGPT in November 2022. *Orthopaedics and Traumatology: Surgery and Research*, *109*(8), 103694. <https://doi.org/10.1016/j.otsr.2023.103694>

Cai, S., & Cui, W. (2023). *Evade ChatGPT detectors via a single space*. <https://arxiv.org/pdf/2307.02599.pdf>

Chaka, C. (2020). *Higher education institutions and the use of online instruction and online tools and resources during the COVID-19 outbreak - An online review of selected U.S. and SA's universities*. <https://doi.org/10.21203/rs.3.rs-61482/v1>

Chaka, C. (2022). Is Education 4.0 a sufficient innovative, and disruptive educational trend to promote sustainable open education for higher education institutions? A review of literature trends. *Frontiers in Education*, *7*(824976), 1-13. <https://doi.org/10.3389/feduc.2022.824976>

Chaka, C. (2023a). Generative AI chatbots - ChatGPT versus YouChat versus Chatsonic: Use cases of selected areas of applied English language studies. *International Journal of Learning, Teaching and Educational Research*, *22*(6), 1-19. <https://doi.org/10.26803/ijlter.22.6.1>

Chaka, C. (2023b). Stylised-facts view of fourth industrial revolution technologies impacting digital learning and workplace environments: ChatGPT and critical reflections. *Frontiers in Education*, *8*, 1150499, 1-10. <https://doi.org/10.3389/feduc.2023.1150499>

Chaka, C. (2023c). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning & Teaching*, *6*(2), 94-104. <https://doi.org/10.37074/jalt.2023.6.2.12>

Chaka, C. (2023d). Fourth industrial revolution—A review of applications, prospects, and challenges for artificial intelligence, robotics and blockchain in higher education. *Research and Practice in Technology Enhanced Learning*, *18*(2), 1-39. <https://doi.org/10.58459/rptel.2023.18002>

Chaka, C. (2023e). The geopolitics of knowledge production in applied English language studies: Transknowledging and a two-eyed critical southern decoloniality. *Journal of Contemporary Issues in Education*, *18*(1), 3-20. <https://doi.org/10.20355/jcie29507>

Cingillioglu, I. (2023). Detecting AI-generated essays: The ChatGPT challenge. *The International Journal of Information and Learning Technology*, *40*(3), 259-268. <https://doi.org/10.1108/IJILT-03-2023-0043>

Cohen, J. (1960). A Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. <https://doi.org/10.1177/001316446002000104>

Cotton, D. R. E., Cotton, P. A., & Shipway, L. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *60*, 1-13. <https://doi.org/10.1080/14703297.2023.2190148>

Dalalah, D., & Dalalah, O. M. A. (2023). The false positives and false negatives of generative AI detection tools in education and academic research: The case of ChatGPT. *The International Journal of Management Education*, *21*(100822),

1-13. <https://doi.org/10.1016/j.ijme.2023.100822>

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, *40*(2), 615-622. <https://doi.org/10.5114/biolport.2023.125623>

Desaire, H. A., Chua, A. E., Isom, M., Jarosova, R., & Hua, D. (2023). Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools. *Cell Reports Physical Science*, *4*(6), 1-2. <https://doi.org/10.1016/j.xcrp.2023.101426>

Domínguez, G. E., Ramírez-March, A., & Montenegro, M. (2023). The geopolitics of knowledge production, or how to inhabit a contradiction: Introduction to the special issue on the narrative productions methodology. *Qualitative Research in Psychology*, *20*(40), 525-541. <https://doi.org/10.1080/14780887.2023.2255104>

Dreibelbis, E. (2023). *OpenAI quietly shuts down AI text-detection tool over inaccuracies*. <https://www.pcmag.com/news/openai-quietly-shuts-down-ai-text-detection-tool-over-inaccuracies>

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 1-63. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>

Elali, F. R., & Rachid, L. N. (2023). AI-generated research paper fabrication and plagiarism in the scientific community. *Patterns*, *4*, 1-4. <https://doi.org/10.1016/j.patter.2023.100706>

Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. *International Journal for Educational Integrity*, *19*(17), 1-16. <https://doi.org/10.1007/s40979-023-00140-5>

Finnegan, M. (2023). *M365 Copilot, Microsoft's generative AI tool, explained*. <https://www.computerworld.com/article/3700709/m365-copilot-microsofts-generative-ai-tool-explained.html>

Gamage, K. A. A., Dehideniya, S. C. P., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: More opportunities than concerns? *Journal of Applied Learning and Teaching*, *6*(2), 358-369. <https://doi.org/10.37074/jalt.2023.6.2.32>

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *Npj Digital Medicine*, *6*(75), 1-5. <https://doi.org/10.1038/s41746-023-00819-6>

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies.

*Health Information and Libraries Journal*, *26*, 91-108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>

Habibzadeh, F. (2023). GPTZero performance in identifying artificial intelligence-generated medical texts: A preliminary study. *Journal of Korean Medical Sciences*, *38*(38), e319. <https://doi.org/10.3346/jkms.2023.38.e319>

Hew, K. F., Huang, W., Du, J., & Jia, C. (2023). Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions. *Journal of Computing in Higher Education*, *35*(1), 40-68. <https://doi.org/10.1007/s12528-022-09338-x>

Homolak, J. (2023). Exploring the adoption of ChatGPT in academic publishing: Insights and lessons for scientific writing. *Croatian Medical Journal*, *64*, 205-207. <https://doi.org/10.3325/cmj.2023.64.205>

Ibrahim, K. (2023). Using AI-based detectors to control AI-assisted plagiarism in ESL writing: "The terminator versus the machines". *Language Testing in Asia*, *13*(46), 1-28. <https://doi.org/10.1186/s40468-023-00260-2>

Ifelebuegu, A. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning and Teaching*, *6*(2), 385-392. <https://doi.org/10.37074/jalt.2023.6.2.2>

Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIED) tools: The good, the bad, and the ugly. *Journal of Applied Learning & Teaching*, *6*(2), 332-345. <https://doi.org/10.37074/jalt.2023.6.2.29>

Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning & Teaching*, *6*(2), 56-63. <https://doi.org/10.37074/jalt.2023.6.2.34>

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, *28*, 15873-15892. <https://doi.org/10.1007/s10639-023-11834-1>

Kasneci, E., SeBler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>

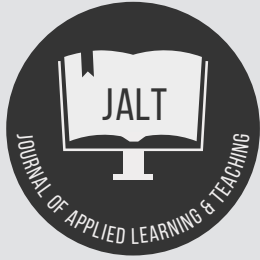
Kenwright, B. (2024). Is it the end of undergraduate dissertations?: Exploring the advantages and challenges of generative AI models in education. In S. Hai-Jew (Ed.), *Generative AI in teaching and learning* (pp. 46-65). IGI Global. <https://doi.org/10.4018/979-8-3693-0074-9.ch003>

Kitchenham, B., & Brereton, P. (2013). A systematic review of systematic review process research in software engineering. *Information and Software Technology*, *55*, 2049-2075. <http://dx.doi.org/10.1016/j.infsof.2013.07.010>

- Kitchenham, B., Brereton, P. O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology, 51*, 7-15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Kleebayoon, A., & Wiwanitkit, V. (2023). Artificial intelligence, chatbots, plagiarism and basic honesty: Comment. *Cellular and Molecular Bioengineering, 16*(2), 173-174. <https://doi.org/10.1007/s12195-023-00759-x>
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability, 15*(7), 5614. <https://doi.org/10.3390/su15075614>
- Krishna, K., Song, Y., Karpinska, M., Wieting, J., & Iyyer, M. (2023). *Paraphrasing evades detectors of AI-generated text, but retrieval is an effective defense*. <https://www.semanticscholar.org/reader/1c13af186d1e177b85ef1ec3fc7b8d33ec314cfd>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies, 28*(1), 973-1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Ladha, N., Yadav, K., & Rathore, P. (2023). AI-generated content detectors: Boon or bane for scientific writing. *Indian Journal of Science and Technology, 16*(39), 3435-3439. <https://doi.org/10.17485/IJST/v16i39.1632>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-74.
- Li, L., Asemota, I., Liu, B., Gomez-Valencia, J., Lin, L., Arif, A. W., ... Usman, M. S. (2022). AMSTAR 2 appraisal of systematic reviews and meta-analyses in the field of heart failure from high-impact journals. *BioMed Central (BMC), 11*(147), 1-8. <https://doi.org/10.1186/s13643-022-02029-9>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), 100790, 1-13. <https://doi.org/10.1016/j.ijme.2023.100790>
- Maddugoda, C. (2023). *A comprehensive review: Detection techniques for human-generated and AI-generated texts*. [https://www.researchgate.net/publication/374542625\\_A\\_Comprehensive\\_Review\\_Detection\\_Techniques\\_for\\_Human-Generated\\_and\\_AI-Generated\\_Texts](https://www.researchgate.net/publication/374542625_A_Comprehensive_Review_Detection_Techniques_for_Human-Generated_and_AI-Generated_Texts)
- McHugh, M. L. (2012). Interrater reliability: The Kappa statistic. *Biochemia Medica, 22*(3), 276-282.
- Müller, M. (2021). Worlding geography: From linguistic privilege to decolonial anywhere. *Progress in Human Geography, 45*(6), 1440-1466. <https://doi.org/10.1177/0309132520979356>
- Nah, F. F.-H., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research, 25*(3), 277-304. <https://doi.org/10.1080/15228053.2023.2233814>
- Odri, G. A., & Yoon, D. J. Y. (2023). Detecting generative artificial intelligence in scientific articles: Evasion techniques and implications for scientific integrity. *Orthopaedics & Traumatology: Surgery & Research, 109*(8), 103706. <https://doi.org/10.1016/j.otsr.2023.103706>
- Perera, P., & Lankathilaka, M. (2023). AI in higher education: A literature review of ChatGPT and guidelines for responsible implementation. *International Journal of Research and Innovation in Social Science (IJRISS), vii*(vi), 306-314. <https://doi.org/10.47772/IJRISS>
- Pérez, J., Díaz, J., Garcia-Martin, J., & Tabuenca, B. (2020). Systematic literature reviews in software engineering—Enhancement of the study selection process using Cohen's Kappa statistic. *The Journal of Systems & Software, 168*, 110657, 1-12. <https://doi.org/10.1016/j.jss.2020.110657>
- Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2023). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics, 1*-25. <https://doi.org/10.1007/s10805-023-09492-6>
- Pinzolit, R. F. J. (2023). AI in academia: An overview of selected tools and their areas of application. *MAP Education and Humanities, 4*, 37-50. <https://doi.org/10.53880/2744-2373.2023.4.37>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching, 6*(1), 41-56. <https://journals.sfu.ca/jalt/index.php/jalt/article/view/787>
- R'boul, H. (2022). Epistemological plurality in intercultural communication knowledge. *Journal of Multicultural Discourses, 17*(2), 173-188. <https://doi.org/10.1080/17447143.2022.2069784>
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems, 3*, 121-154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching, 6*(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Santra, P. P., & Majhi, D. (2023). Scholarly communication and machine-generated text: Is it finally AI vs AI in plagiarism detection? *Journal of Information and Knowledge, 60*(3), 175-183. <https://doi.org/10.17821/srels/2023/v60i3/171028>
- Sarzaeim, P., Doshi, A. M., & Mahmoud, Q. H. (2023). A

- framework for detecting AI-generated text in research publications. <https://proceedings.icatsconf.org/conf/index.php/ICAT/article/download/36/21/117>
- Sayed, A. M-U. (2023). *Reliably detecting AI-generated text is mathematically impossible*. <https://www.linkedin.com/pulse/reliably-detecting-ai-generated-text-mathematically-sayed>
- Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., ... Boers, M. (2009). AMSTAR is a reliable and valid measurement to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology*, 62, 1013-1020. <https://doi.org/10.1016/j.jclinepi.2008.10.009>
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., ... Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *British Medical Journal*, 358, 1-9. <http://dx.doi.org/10.1136/bmj.j4008>
- Singh, A. (2023). *A comparison study on AI language detector*. <https://doi.org/10.1109/CCWC57344.2023.10099219>.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Stracke, C. M., Chounta, I-A., Holmes, W., Tlili, A., & Bozkurt, A. (2023). A standardised PRISMA-based protocol for systematic reviews of the scientific literature on Artificial Intelligence and Education (AI&ED). *Journal of Applied Learning and Teaching*, 6(2), 64-70. <https://doi.org/10.37074/jalt.2023.6.2.38>
- Subramaniam, R. (2023). Identifying text classification failures in multilingual AI-generated content. *International Journal of Artificial Intelligence and Applications (IJAA)*, 14(5), 57-63. <https://doi.org/10.5121/ijaia.2023.14505>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 31-40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Sumakul, D. T. Y. G., Hamied, F. A., & Sukyadi, D. (2022). Artificial intelligence in EFL classrooms: Friend or foe? *LEARN Journal: Language Education and Acquisition Research Network*, 15(1), 232-256.
- Thurzo, A., Strunga, M., Urban, R., Surovková, J., & Afrashtehfar, K. I. (2023). Impact of artificial intelligence on dental education: A review and guide for curriculum. *Education Sciences*, 13(150), 1-15. <https://doi.org/10.3390/educsci13020150>
- Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education & Technology*, 3(1), 45-54.
- van Dis, E. A., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224-226.
- Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated writing: A comparison of 16 AI text detectors. *Open Information Science*, 7(20220158), 1-24.
- Wang, Y., Mansurov, J., Ivanov, P., Su, J., Shelmanov, A., Tsvigun, A., ... Nakov, P. (2023). *M4: Multi-generator, multi-domain, and multi-lingual black-box machine-generated text detection*. <https://arxiv.org/pdf/2305.14902.pdf>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... Waddington, L. (2023). *Testing of detection tools for AI-generated text*. <https://doi.org/10.48550/arXiv.2306.15666>
- Wee, H. B., & Reimer, J. D. (2023). Non-English academics face inequality via AI-generated essays and countermeasure tools. *BioScience*, 73, 476-478. <https://doi.org/10.1093/biosci/biad034>
- Wohlin, C., Kalinowski, M., Felizardo, K. R., & Mendes, E. (2022). Successful combination of database search and snowballing for identification of primary studies in systematic literature studies. *Information and Software Technology*, 147(106908), 1-12. <https://doi.org/10.1016/j.infsof.2022.106908>
- Wu, J., Yang, S., Zhan, R., Yuan, Y., Wong, D. F., & Chao, L. S. (2023). *A survey on LLM-generated text detection: Necessity, methods, and future directions*. <https://arxiv.org/pdf/2310.14724.pdf>
- Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39(1), 93-112. <https://doi.org/10.1177/0739456X17723971>
- Yang, J., Chen, Y. L., Por, L. Y., & Ku, C. S. (2023). A systematic literature review of information security in chatbots. *Applied Sciences*, 13(11), 6355. <https://doi.org/10.3390/app13116355>
- Yeadon, W., Inyang, O-O., Mizouri, A., Peach, A., & Testrow, C. P. (2023). The death of the short-form physics essay in the coming AI revolution. *Physics Education*, 58(035027), 1-13.

Copyright: © 2024. Chaka Chaka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Accuracy pecking order – How 30 AI detectors stack up in detecting generative artificial intelligence content in university English L1 and English L2 student essays

---

Chaka Chaka<sup>A</sup>

A

*Professor, University of South Africa, Pretoria, South Africa*

---

### Keywords

Accuracy;  
AI;  
AI detectors;  
AI-generated and human-written content;  
artificial intelligence;  
English L1 and English L2;  
false positive rates;  
student essays;  
true negative rates.

### Abstract

This study set out to evaluate the accuracy of 30 AI detectors in identifying generative artificial intelligence (GenAI)-generated and human-written content in university English L1 and English L2 student essays. 40 student essays were divided into four essay sets of English L1 and English L2 and two undergraduate modules: a second-year module and a third-year module. There are ten essays in each essay set. The 30 AI detectors comprised freely available detectors and non-premium versions of online AI detectors. Employing a critical studies approach to artificial intelligence, the study had three research questions. It focused on and calculated the accuracy, false positive rates (FPRs), and true negative rates (TNRs) of all 30 AI detectors for all essays in each of the four sets to determine the accuracy of each AI detector to identify the GenAI content of each essay. It also used confusion matrices to determine the specificity of best- and worst-performing AI detectors. Some of the results of this study are worth mentioning. Firstly, only two AI detectors, Copyleaks and Undetectable AI, managed to correctly detect all of the essay sets of the two English language categories (English L1 and English L2) as human written. As a result, these two AI detectors jointly shared the first spot in terms of the GenAI detection accuracy ranking. Secondly, nine of the 30 AI detectors completely misidentified all the essays in each of the four essay sets of the two language categories in both modules. Thus, they collectively shared the last spot. Thirdly, the remaining 19 AI detectors both correctly and incorrectly classified the four essay sets in varying degrees without any bias to any essay set of the two English language categories. Fourthly, none of the 30 AI detectors tended to have a bias toward a specific English language category in classifying the four essay sets. Lastly, the results of the current study suggest that the bulk of the currently available AI detectors, especially the currently available free-to-use AI detectors, are not fit for purpose.

---

### Correspondence

chakachaka8@gmail.com <sup>A</sup>

---

### Article Info

Received 9 March 2024

Received in revised form 28 March 2024

Accepted 3 April 2024

Available online 10 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.33>

## Introduction

In academia, plagiarism and generative artificial intelligence (GenAI)-generated content are two different things. For instance, a student does not need a GenAI tool to plagiarise, but they need a GenAI tool to generate GenAI content. Notably, plagiarism predates the advent of GenAI content generation, especially as the latter is heralded by GenAI language models such as ChatGPT. As such, the possibility of plagiarism is always there with or without the use of GenAI tools, but GenAI-generated content is almost impossible to generate without using GenAI tools such as ChatGPT as its catalysts. With the launch of ChatGPT and the other related GenAI-powered chatbots, the quest for detecting GenAI-generated content in university student writing, in particular, has become unavoidable. What is even more pressing is the quest for differentiating between GenAI-generated and human-written content in student writing in higher education (HE). In the HE arena, universities and academics have always prided themselves in being the guardians and protectors of original and authentic academic writing in all disciplines. This guardianship and protectorship has often come under the banner of academic integrity (see Anthology White Paper, 2023; Blau et al., 2020; Gamage et al., 2020; Perkins, 2023; Sullivan et al., 2023; Uzun, 2023). It is no exaggeration to assert that academic integrity, guardianship and protectorship in HE almost borders on a frenzy due to, mainly, though not exclusively, pressure points brought by GenAI-powered chatbots like ChatGPT. In this frenzied scrambling, GenAI-generated content and plagiarism feature as proxies for academic dishonesty.

However, viewing academic integrity through the prism of its nemesis, like academic dishonesty that comprises GenAI-generated content and plagiarism, is simplistic and superficial. This conception of academic integrity has to do with the practice of text- or content-matching that chimes with plagiarism-detection software programmes in which plagiarism and GenAI-generated content, are deemed a twin threat to academic integrity (cf. Blau et al., 2020; Gamage et al., 2020; Ifelebuegu, 2023; Rudolph et al., 2023; Sobaih, 2024). As Gamage et al. (2020) contend, this view of academic integrity overlooks other elements of academic dishonesty or other violations of academic integrity (see Blau et al., 2020). In addition to GenAI-generated content and plagiarism, examples of elements of academic dishonesty or violations of academic integrity include fraudulence, falsification, fabrication, facilitation, cheating, ghost-writing (Blau et al., 2020), contract cheating, and collusion (Gamage et al., 2020). Of course, some of these elements or violations may overlap: fraudulence with falsification and fabrication, ghost-writing with contract cheating, and facilitation with collusion (cf. Blau et al., 2020; Gamage et al., 2020). Additionally, both cheating and fraudulence can be used as overarching terms for academic dishonesty. Therefore, reducing academic dishonesty to GenAI-generated content and plagiarism alone tends to obscure its other facets, such as the ones furnished here.

With the surge of GenAI-generated content and plagiarism being a threat to academic integrity in HE, several AI content detectors have been released, while existing traditional plagiarism detection tools have upgraded their offerings to

include AI content detection features (see Anil et al., 2023; Chaka, 2023a, 2024; Bisi et al., 2023; Dergaa et al., 2023; Ladha et al., 2023; Uzun, 2023; Wiggers, 2023; Weber-Wulff et al., 2023). The cardinal function of AI content detectors is to do exactly what they are designed to do: detect GenAI-generated content in different types of academic and scholarly writing. To this effect, there have been studies that have tested the effectiveness or reliability of AI content detectors in detecting GenAI-generated content in academic writing, or in distinguishing between GenAI-generated and human-written content in academic writing. These studies have tested different types of AI content detectors that include single AI content detectors (see Habibzadeh, 2023; Perkins et al., 2024; Subramaniam, 2023), two AI content detectors (see Bisi et al., 2023; Desaire et al., 2023; Ibrahim, 2023), three AI content detectors (see Cingillioglu, 2023; Elali & Rachid, 2023; Gao et al., 2023; Homolak, 2023; Ladha et al., 2023; Wee & Reimer, 2023), four AI content detectors (Abani et al., 2023; Alexander et al., 2023; Anil et al., 2023), and multiple AI content detectors (Chaka, 2023a; Odri & Yoon, 2023; Santra & Majhi, 2023; Walters, 2023) (see Chaka, 2024).

Most crucially, there is one study that has discovered that AI detectors tend to be biased against non-English language speakers (Liang et al., 2023; Mathewson, 2023; Shane, 2023; cf. Adamson, 2023; Gillham, 2024). This finding resonates, in a different but related scenario, with the view that some studies have established that currently available automatic speech recognition technologies poorly detect, if any, and discriminate against the English spoken by Black people, especially African American Language (AAL), thereby exposing their racial bias and demographic discrimination against this type of English (Martin & Wright, 2023). Linguistic and racial biases are but two of the instances of bias that GenAI models, and not just AI detection models, have to contend with in their everyday deployment. Other instances of bias GenAI models have to grapple with are cultural, ideological, political, temporal, and confirmation biases (see Ferrara, 2023). Thus, in addition to simply detecting GenAI-generated content, or distinguishing it from its human-written counterpart, these biases are some of the pressing challenges that these models have to wrestle with on an ongoing basis.

Against this background, the current study set out to:

- evaluate the accuracy of 30 AI detectors in differentiating between GenAI-generated and human-written content in university English L1 and English L2 student essays for two different undergraduate modules;
- establish whether these 30 AI detectors will classify these four sets of student essays differentially based on their English L1 and English L2 categories; and
- discover which language category within these four sets of student essays is assigned more false positives.



On this basis, the overarching purpose of this study is to contribute to the ongoing debate about the effectiveness (accuracy, precision, and reliability) of AI content detectors in distinguishing between GenAI-generated and human-written content in the essays produced by English L1 and English L2 students. The student essays in this study were written by English L1 and English L2 students who registered for a second-year undergraduate module and a third-year undergraduate module offered by an English department at a university in South Africa in 2018, 2020, and 2022.

Given the points highlighted above, this study seeks to answer the following research questions (RQs):

- RQ1: What is the accuracy of the 30 AI detectors in differentiating between GenAI-generated and human-written content in university English L1 and English L2 student essays for two different undergraduate modules?
- RQ2: Do these 30 AI detectors classify these four sets of student essays differentially based on their English L1 and English L2 categories or not?
- RQ3: Which language category within these four sets of student essays is assigned more false positives by these AI detectors?

### Critical studies approach to AI

In a surreal world, AI, algorithms, and machine learning would be devoid of any bias: racial, demographic, gender, sexuality, disability, and training data bias (see Lindgren, 2023; also see AIContentfy team, 2023; Chaka, 2022; Ferrara, 2023; Wu et al., 2023). In real-world contexts, though, that is not the case. This rings true for AI detectors. Their efficacy is largely determined by, among other things, their training data, their algorithms, and their computing prowess (AIContentfy team, 2023). All of this, together with the types of bias mentioned and those stated earlier, leads to AI detectors having shortcomings and deficiencies. As such, they end up not being as effective and efficient as they are made out to be or as they often claim to be. This is where a critical studies approach to AI comes in. This approach draws on some of the ideas propounded by Chaka (2022), Couldry and Mejias (2019), Lindgren (2023), Mohamed et al. (2020), Ricaurte (2019), who adopt a critically driven approach to dealing with and studying technology, algorithms, data, and datafication. Importantly, it draws on Lindgren's (2023) notion of critical studies of AI.

In this paper, in particular, the critical studies approach to AI entails recognising that AI detectors are not 100% efficient and effective: they have limitations, deficiencies, and biases. This is so notwithstanding the accuracy percentage claims that these models may arrogate to themselves on their landing pages. This approach also acknowledges that AI detectors are constrained by contextual factors such as domains, algorithms, training data, performance, robustness, and adversarial testing. The latter refers to how well an AI detector performs when tested with an adversarial input like edited or paraphrased content (see Captain Words, 2024; Wu et al., 2023) or such as single spacing (Cai & Cui, 2023).

This latter aspect highlights the fact that AI detectors can be tricked by manipulating or reworking input content (see Chaka, 2023a; Lee, 2023). This is one of the limitations AI detectors have, which is recognised by the critical studies approach to AI as framed here. Finally, this approach contends that the limitations and deficiencies of AI detectors should not be reduced to technologism alone: they are also a reflection of their designers, architecture, or otherwise.

### Related literature

This related literature section is unconventional in that it selectively deals with a few studies that have a bearing on the current study. To this end, it wants to foreground a few points. First, save for Liang et al.'s (2023) study, there is a paucity of studies that have tested how currently available AI detectors tend to be biased against non-native English writers/students vis-à-vis native English writers/students. Secondly, as pointed out briefly earlier, since the release of ChatGPT and the other related GenAI-powered chatbots, several AI detectors have been designed and launched, which are intended to detect GenAI-generated content or distinguish between GenAI-generated and human-written content. In keeping with this attempt to detect GenAI-generated content, existing traditional plagiarism detection software programmes have been upgraded to accommodate AI detection tools in their offerings (see Anil et al., 2023; Bisi et al., 2023; Chaka, 2023a, 2024; Dergaa et al., 2023; Ladha et al., 2023; Uzun, 2023; Wiggers, 2023; Weber-Wulff et al., 2023). Again, as stated earlier, some studies have evaluated the effectiveness of single AI detectors (Habibzadeh, 2023; Subramaniam, 2023), two AI detectors (Desaire et al., 2023; Ibrahim, 2023), three AI detectors (Cingillioglu, 2023; Elali & Rachid, 2023; Wee & Reimer, 2023), four AI detectors (Alexander et al., 2023; Anil et al., 2023), and multiple AI detectors (Chaka, 2023a, 2024; Odri & Yoon, 2023; Walters, 2023).

In the midst of so many and varied studies that have been conducted in the aftermath of ChatGPT's launch, I will, in this section, briefly discuss a select few studies that have explored or tested the effectiveness of multiple AI detectors in detecting GenAI-generated content or distinguish between GenAI-generated from human-written content in given subject areas. Elsewhere, Chaka (2024) conducted a review of studies that tested the effectiveness of different AI detectors in distinguishing between GenAI-generated and human-written content in different subject areas. It is also worth mentioning that some of the studies that have investigated the effectiveness of multiple AI detectors, in this regard, are preprints like Webber-Wulff (2023) and Wu et al. (2023). Others are AI detectors' in-house studies such as AIContentfy Team, 2023; Captain Words, 2024). The first study that has some bearing on the present study is Liang et al.'s (2023) study. This study set out to evaluate the effectiveness of seven AI detectors in detecting GenAI-generated text in a dataset of 91 human-written Test of English as a Foreign Language (TOEFL) essays and in a dataset of 88 U.S. 8th-grade essays extracted from the Hewlett Foundations' Automated Student Assessment Prize (ASAP). The first dataset was sourced from a Chinese educational forum. The seven AI detectors employed to evaluate these

two essay datasets were ZeroGPT, GPTZero, Crossplag, OpenAI, Sapling, Quillbot, and Originality. These detectors detected and classified the U.S. 8th-grade essay dataset almost accurately. Nonetheless, they misidentified more than half of the TOEFL essay dataset as generated by GenAI, with a mean false positive rate (FPR) of 61.22%. In addition, these AI detectors accorded the misidentified TOEFL essays a very low perplexity due to the limited linguistic variability of these essays, which was easily predictable. But, after ChatGPT was employed to improve the linguistic expressions of the TOEFL essays to those of a native English speaker, their misidentification by the said AI detectors decreased, with their mean FPR concomitantly decreasing to 11.77%, and their perplexity significantly improving as well.

Since the publication of Liang et al.'s (2023) study, there have been, in varying degrees, some comments about it (see Mathewson, 2023; Shane, 2023) and some reactions to it (see Adamson, 2023; Gillham, 2024). Among the reactions, Adamson's (2023) is the most interesting one as it shows how Liang et al.'s (2023) study seems to have ruffled up the veneer of AI detectors' effectiveness in detecting GenAI-generated text in student-written essays without being linguistically biased. To this effect, a Turnitin test was subsequently conducted to detect GenAI-generated text in three datasets of ASAP, ICNALE, and PELIC that comprised L1 English (ASAP = 2,481 and ICNALE = 400) and L2 English (ICNALE = 2,222 and PELIC = 4,000). The results of this test showed that for documents with a minimum 300-word threshold, the difference in the false positive rate (FPR) between L1 English essays and L2 English essays was fractional and, thus, was not statistically significant. This proved that the paper asserts that Turnitin, as an AI detector, did not evince any statistically significant bias against the two sets of English language essays. Moreover, the paper avers that even though each essay set's FPR was marginally higher than Turnitin's overall target of 0.01 (1%), none of the two essay sets' FPR was significantly different from this overall target. In contrast, the paper argues that for documents whose content was below the minimum 300-word threshold, there was a significant difference in the FPR between L1 English essays and L2 English essays. This difference was greater than Turnitin's 0.01 overall target. On this basis, the paper concludes that this finding confirms that AI detectors need longer essay samples for them to detect GenAI-generated content accurately and for them to be able to avoid producing a high rate of false positives (Adamson, 2023). An overall FPR target of 1% means that 10 human-produced student essays are likely to be misclassified as false positives in every 1,000 university essay scripts. This number is still concerning given those students who might be affected by this misclassification (see Anderson, 2023).

It is worth mentioning that Turnitin is not among the seven AI detectors tested by Liang et al. (2023). Despite this, there is no gainsaying that this resultant Turnitin test bears testimony to the ruffle that Liang et al.'s (2023) study has caused to the AI detection ecosystem, not only Turnitin but that of the other AI detectors as well. The other point to emphasise is that Liang et al.'s (2023) study has an element of a critical studies approach to AI. This element has to do with the way the study approached the seven AI detectors from a critical standpoint by highlighting their

linguistic detection bias in dealing with native English speakers versus non-native English speakers in their written English. Moreover, this criticality element is related to the two adversarial prompts the study inputted into ChatGPT to write the two datasets differently with a view of tricking the seven AI detectors. It is when one applies this type of critical perspective which is grounded on relevant raw data to GenAI in general, and to AI detectors in particular, that one gets the owners and designers of AI detectors' attention as is the case with Adamson's (2023) paper. Without that criticality, nothing is likely to happen.

Among the studies that have evaluated multiple AI detectors in other subject areas than English is Odri and Yoon's (2023) study. This study had three objectives, which were to: evaluate 11 AI detectors' performance on a wholly GenAI-generated text, test AI detection-evading methods, and evaluate how effective these AI detection-evading methods were on previously tested AI detectors. It hypothesised that the 11 AI detectors to be tested were not all equally effective in identifying GenAI-generated text and that some of the evasion methods could render the GenAI-generated text almost undetectable. The GenAI text was generated from ChatGPT-4 and was tested on 11 AI detectors: Originality, ZeroGPT, Writer, Copyleaks, Crossplag, GPTZero, Sapling, Content at Scale, Corrector, Writfull, and Quill. The text was tested before applying AI detection evasion techniques and after applying them. The AI detection evasion techniques employed included: improving command messages (prompts) in ChatGPT, adding minor grammatical errors (e.g., a comma deletion), paraphrasing, and substituting Latin letters with their Cyrillic equivalents. The GenAI text was manipulated six times to produce its slightly modified versions using the aforesaid evasion techniques in ChatGPT. The study also tested a scientific text produced by a human (Sir John Charnley) in 1960 (Odri & Yoon, 2023). One plausible reason that can be extrapolated from the study about the use of this text is that it is freely available online. The other plausible reason is that the text predates the advent of GenAI models, particularly ChatGPT, by 62 years. Therefore, in 1960, there was no way any text could have been generated by GenAI models.

For the initial, unaltered GenAI text generated by ChatGPT, seven of the 11 AI detectors identified it as written mainly by humans. This is how these AI detectors fared in this text: GPTZero = human, Writer = 100% human, Quill = human, Content at Scale = 85% human, Copyleaks = 59.9% human, Corrector = 0.02% AI, and ZeroGPT = 25.8% AI. The more this text was slightly modified in sustained degrees (one modification after another as mentioned above), the more the 11 AI detectors misclassified it as human-written. Regarding the human-written text, only one of the 11 AI detectors (Originality) was able to correctly detect it as having 0% AI. It is important to mention that despite this correct detection, Originality is one of the four AI detectors that misidentified the final modified version of the GenAI-generated text as having 0% AI content (Odri & Yoon, 2023). Like Liang et al.'s (2023) study discussed above, the relevance of Odri and Yoon's (2023) study is that it has elements of a critical studies approach to AI. Its use of adversarial attacks in the form of prompt attacks is an example of an adversarial input that I earlier referred to as one of the contextual factors that

degrades the efficacy of AI detectors (also see Anderson, 2023; Chaka, 2023a, 2024; Krishna et al., 2023; Sadasivan et al., 2023). From a critical perspective, prompt attacks expose the limitations and deficiencies of AI detectors.

## Materials and methods

This study followed an exploratory research design, with the primary objective of exploring a given area, aspect, or phenomenon that has not been extensively researched. By its nature, exploratory research can tentatively analyse a new emerging topic, or suggest new ideas (Swedberg, 2020; see Makri & Neely, 2021). Testing the accuracy and effectiveness of AI detectors in identifying GenAI-generated and human-written content, or in distinguishing between these content types is still a relatively new area in many disciplines (see Chaka, 2023a, 2023b).

## Data collection

The data collection process for this study comprised three stages. The first stage entailed selecting student (human) essay samples. These essays consisted of four datasets of university English L1 and English L2 student essays. They were selected from a pool of essays that had been submitted as assignment responses for two undergraduate modules offered by an English department at an open-distance and e-learning university in South Africa. The modules were second and third-year, major modules. Each dataset had ten essays. The two sets of essays for a second-year major module were submitted in 2018 (second semester), 2020 (first and second semesters), and 2022 (first and second semesters). The submission details of the ten essays in the English L1 essay set were as follows: 2018 first semester (n = 1), 2020 first semester (n = 4), 2020 second semester (n = 3), 2022 first semester (n = 1), and 2022 second semester (n = 1). The corresponding English L2 essay set for the second-year module consisted of the following essays in relation to their years and semesters of submission: 2020 first semester (n = 3), 2022 first semester (n = 1), and 2022 second semester (n = 6). Both sets of essays (English L1 and English L2) for a third-year, major module, each of which with ten essays, were submitted in the first semester of 2020.

As is evident from the points presented above, the four datasets used in this study together had 40 essays. The essays were randomly selected from assignment scripts that served as either dummy or moderation scripts that are generally emailed to module team members by module primary lecturers. It is from this pool of essays that the current student essays were selected for this study. These essays were categorised as English L1 and English L2 based on whether the students who wrote them had identified English as their home language (English L1) or had identified a different language other than English as their home language (English L2) in their module registration information. All the selected essays for the four datasets were copied from their original PDF files and pasted into an MS Word file without changing anything. Thereafter, two MS Word files, English L1 and English L2 essay sets, were

compiled for the two modules. The ten English L1 essays for the second-year module had a total word count of 4,465, with a mean word count of 446.5; their counterpart English L2 essays had a total word count of 4,322, with a mean word count of 432.2. The total word count of the ten English L1 essays for the third-year module was 4,504, with a mean word count of 450.4. Their corresponding English L2 essays had a total word count of 4,404, with 440.4 as their mean word count. The essay selection and compiling process took place between 18 December 2023 and 20 December 2023. Before the study was conducted, ethical clearance was secured, and the certificate number of this ethical clearance is Ref #: 2021\_RPSC\_050.

The second stage in the data collection process involved choosing free, publicly available online AI detectors. This process happened between 21 December 2023 and 28 December 2023. During which, many online AI detectors were identified. After trialling some of them, 30 AI detectors were chosen for use in this study (see Table 1). Then, from 02 January 2024 to 20 February 2024, the third stage occurred. Each essay from the four datasets was submitted to each of the 30 AI detectors for GenAI-generated content scanning. The test scores for each essay scan were copied and transferred to relevant tables, each of which was labelled English L1 and English L2 for each of the two modules, with each AI detector's name used as a caption for each table. However, to avoid having 30 individual tables, two tables were merged into one (see Table 1).

Table 1: Names of 30 AI detectors and their accuracy ranking.

Rank No.	Names of AI detectors	Rank No.	Names of AI detectors
1.	<b>Copyleaks</b> 2nd year: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) 3rd-year module: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) Score for Accuracy and TNR = 8	9.	<b>Rank Wizard AI</b> 2nd year: L1 (FPR = 0.4, Accuracy = 0.6, TNR = 0.6), L2 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8) 3rd-year module: L1 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4), L2 (FPR = 0.4, Accuracy = 0.6, TNR = 0.6) Score for Accuracy and TNR = 4.8
2.	<b>Undetectable AI</b> 2nd year: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) 3rd-year module: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) Score for Accuracy and TNR = 8	10.	<b>GTZero</b> 2nd year: L1 (FPR = 0.8, Accuracy = 0.2, TNR = 0.2), L2 (FPR = 0.5, Accuracy = 0.5, TNR = 0.5) 3rd-year module: L1 (FPR = 0.5, Accuracy = 0.5, TNR = 0.5), L2 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4) Score for Accuracy and TNR = 3.2
3.	<b>Five Moderation</b> 2nd year: L1 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9), L2 (FPR = 0, Accuracy = 1, TNR = 1) 3rd-year module: L1 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9), L2 (FPR = 0, Accuracy = 1, TNR = 1) Score for Accuracy and TNR = 7.6	11.	<b>Corrector App</b> 2nd year: L1 (FPR = 0.7, Accuracy = 0.3, TNR = 0.3), L2 (FPR = 0.7, Accuracy = 0.3, TNR = 0.3) 3rd-year module: L1 (FPR = 1, Accuracy = 0.0, TNR = 0.0), L2 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7) Score for Accuracy and TNR = 2.6
4.	<b>Scribbr</b> 2nd year: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) 3rd-year module: L1 (FPR = 0, Accuracy = 0.9, TNR = 0.9), L2 (FPR = 0, Accuracy = 0.9, TNR = 0.9) Score for Accuracy and TNR = 7.6	12.	<b>SciSpace AI Detector</b> 2nd year: L1 (FPR = 0.7, Accuracy = 0.3, TNR = 0.3), L2 (FPR = 0.4, Accuracy = 0.6, TNR = 0.6) 3rd-year module: L1 (FPR = 1, Accuracy = 0.0, TNR = 0.0), L2 (FPR = 0.7, Accuracy = 0.3, TNR = 0.3) Score for Accuracy and TNR = 2.4
5.	<b>AI Content Detector</b> 2nd year: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) 3rd-year module: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7) Score for Accuracy and TNR = 7.4	13.	<b>Content at Scale</b> 2nd year: L1 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4), L2 (FPR = 0.7, Accuracy = 0.3, TNR = 0.3) 3rd-year module: L1 (FPR = 0.5, Accuracy = 0.5, TNR = 0.5), L2 (FPR = 0.5, Accuracy = 0.5, TNR = 0.5) Score for Accuracy and TNR = 1.6
6.	<b>Plagiarism Detector</b> 2nd year: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9) 3rd-year module: L1 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9), L2 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9) Score for Accuracy and TNR = 7.4	14.	<b>StrahWriter</b> 2nd year: L1 (FPR = 0.7, Accuracy = 0.3, TNR = 0.3), L2 (FPR = 0.8, Accuracy = 0.2, TNR = 0.2) 3rd-year module: L1 (FPR = 0.9, Accuracy = 0.1, TNR = 0.1), L2 (FPR = 0.9, Accuracy = 0.1, TNR = 0.1) Score for Accuracy and TNR = 1.4
7.	<b>Duipi Checker</b> 2nd year: L1 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9), L2 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9) 3rd-year module: L1 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8), L2 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7) Score for Accuracy and TNR = 6.6	15.	<b>QuillBot AI Detector</b> 2nd year: L1 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4), L2 (FPR = 0.9, Accuracy = 0.1, TNR = 0.1) 3rd-year module: L1 (FPR = 0.9, Accuracy = 0.1, TNR = 0.1), L2 (FPR = 1, Accuracy = 0.0, TNR = 0.0) Score for Accuracy and TNR = 1.2
8.	<b>Grammarly</b> 2nd year: L1 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7), L2 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9) 3rd-year module: L1 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9), L2 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8) Score for Accuracy and TNR = 6.6	16.	<b>AI Content Checker</b> 2nd year: L1 (FPR = 1, Accuracy = 0.0, TNR = 0.0), L2 (FPR = 1, Accuracy = 0.0, TNR = 0.0) 3rd-year module: L1 (FPR = 1, Accuracy = 0.0, TNR = 0.0), L2 (FPR = 1, Accuracy = 0.0, TNR = 0.0) Score for Accuracy and TNR = 0.0
9.	<b>ZeroGPT</b> 2nd year: L1 (FPR = 0, Accuracy = 1, TNR = 1), L2 (FPR = 0, Accuracy = 1, TNR = 1) 3rd-year module: L1 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4), L2 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8) Score for Accuracy and TNR = 6.4		<b>AI-Detector</b> 2nd year: L1 (FPR = 1, Accuracy = 0.0, TNR = 0.0), L2 (FPR = 1, Accuracy = 0.0, TNR = 0.0) 3rd-year module: L1 (FPR = 1, Accuracy = 0.0, TNR = 0.0), L2 (FPR = 1, Accuracy = 0.0, TNR = 0.0) Score for Accuracy and TNR = 0.0
10.	<b>Detect Bard</b> 2nd year: L1 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8), L2 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9) 3rd-year module: L1 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4), L2 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8) Score for Accuracy and TNR = 5.8		
11.	<b>AI Checker Tool</b> 2nd year: L1 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8), L2 (FPR = 0.1, Accuracy = 0.9, TNR = 0.9) 3rd-year module: L1 (FPR = 0.6, Accuracy = 0.4, TNR = 0.4), L2 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7) Score for Accuracy and TNR = 5.6		
12.	<b>AI Contently</b> 2nd year: L1 (FPR = 0.4, Accuracy = 0.6, TNR = 0.6), L2 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7) 3rd-year module: L1 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7), L2 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8) Score for Accuracy and TNR = 5.6		
13.	<b>Writer</b> 2nd year: L1 (FPR = 0.2, Accuracy = 0.8, TNR = 0.8), L2 (FPR = 0.3, Accuracy = 0.7, TNR = 0.7) 3rd-year module: L1 (FPR = 0.4, Accuracy = 0.6, TNR = 0.6), L2 (FPR = 0.5, Accuracy = 0.5, TNR = 0.5) Score for Accuracy and TNR = 5.2		

```

AI Detector
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

Detecting AI.com
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

GLTR
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

GPT-2 Output Detector Demo
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

TryPaids GPT Essay Checker
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

ReWriteGuru's AI Detector
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

SEO
2nd year: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
3rd year module: L1 (FPR = 1; Accuracy = 0.0; TNR = 0.0); L2 (FPR = 1; Accuracy = 0.0; TNR = 0.0)
Score for Accuracy and TNR = 0.0

```

## Data analysis

After the scan results for each of the relevantly labelled tables had been captured under the English L1 and English L2 categories for each of the two modules, the GenAI and human content probability scores (as percentages) and their accompanying statements as yielded by each AI detector, were entered in an MS Word file. The GenAI and human content probability scores for each set of English L1 and English L2 essays were calculated and summed. The sum for each set was averaged to get the mean score. This procedure was done for all essay datasets whose AI detector scans yielded GenAI and human content probability scores. For those essay datasets whose AI detector scans yielded only statements, those statements were captured accordingly in a tabular form. The mean scores of all the scan scores for all AI detectors were compared in each language category. Additionally, false positives (human-written essays misclassified as GenAI-generated) and true negatives (correctly detected human-written essays) for each AI detector were calculated with a view to getting false positive rates (FPRs) and true negative rates (TNRs) within each AI detector and between all AI detectors. The accuracy, specificity, and negative predictive value (NPV) of AI detectors whose test results were a direct opposite of each other were measured using confusion matrices (see Captain Words, 2024; Colquhoun, 2014; Gillham, 2024; Weber-Wulff et al., 2023; Wu et al., 2023) and compared with those of its counterparts.

## Results

The GenAI test scores that were yielded by scanning each of the 30 AI detectors were compiled in a table (see Table 2). These test results were captured in the manner in which each AI detector displayed them without any modification. An example of such results is shown in Table 2. The exception is the phrasing about the colour red and the colour purple provided for GLTR AI test results. But even for this AI detector, this phrasing was formulated in keeping with how this AI detector itself explains its colour-coded scan scores. Where each AI detector's scan scores made it possible, the GenAI and human content probability scores for each set of English L1 and English L2 essays, together with their respective means, were calculated (see Tables 2

and 3). As is evident from Table 2, various GenAI and human content probability scores, expressed in percentages and percentage points, have been displayed as generated by Writer's and ZeroGPT's scan scores (raw data) for each of the ten essays for each of the two sets of essays for English L1 and English L2. These two AI detectors are used here for illustrative purposes since the scan scores of each of the 30 AI detectors cannot be displayed for lack of space. For example, Writer detected eight essays and seven essays for L1 and L2, respectively, under the 2nd-year module, as having 100% human-generated content. For the 3rd-year module, Writer classified six essays and five essays for L1 and L2, apiece, as containing 100% human-generated content. In contrast, under the 2nd-year module, ZeroGPT classified nine essays and none as containing 0% AI GPT content for L1 and L2 respectively. It, then, identified four essays for L1 and eight essays for L2 under the 3rd-year module, as having 0% AI GPT content.

Table 2: An example of how scan/test results were captured.

Writer				ZeroGPT			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1	L2	L1	L2	L1	L2	L1	L2
Essay 1 = 100% human-generated content	Essay 1 = 100% human-generated content	Essay 1 = 100% human-generated content	Essay 1 = 100% human-generated content	Essay 1 = 0% AI GPT	Essay 1 = 0% AI GPT	Essay 1 = 14.91% AI GPT	Essay 1 = 19.42% AI GPT
Essay 2 = 100% human-generated content	Essay 2 = 100% human-generated content	Essay 2 = 98% human-generated content	Essay 2 = 100% human-generated content	Essay 2 = 0% AI GPT	Essay 2 = 0% AI GPT	Essay 2 = 5.35% AI GPT	Essay 2 = 0% AI GPT
Essay 3 = 100% human-generated content	Essay 3 = 100% human-generated content	Essay 3 = 100% human-generated content	Essay 3 = 100% human-generated content	Essay 3 = 0% AI GPT	Essay 3 = 0% AI GPT	Essay 3 = 16.49% AI GPT	Essay 3 = 0% AI GPT
Essay 4 = 97% human-generated content	Essay 4 = 100% human-generated content	Essay 4 = 100% human-generated content	Essay 4 = 99% human-generated content	Essay 4 = 6.88% AI GPT	Essay 4 = 0% AI GPT	Essay 4 = 0% AI GPT	Essay 4 = 0% AI GPT
Essay 5 = 100% human-generated content	Essay 5 = 100% human-generated content	Essay 5 = 94% human-generated content	Essay 5 = 99% human-generated content	Essay 5 = 0% AI GPT	Essay 5 = 0% AI GPT	Essay 5 = 0% AI GPT	Essay 5 = 0% AI GPT
Essay 6 = 100% human-generated content	Essay 6 = 100% human-generated content	Essay 6 = 100% human-generated content	Essay 6 = 100% human-generated content	Essay 6 = 0% AI GPT	Essay 6 = 0% AI GPT	Essay 6 = 16.96% AI GPT	Essay 6 = 0% AI GPT
Essay 7 = 100% human-generated content	Essay 7 = 95% human-generated content	Essay 7 = 35% human-generated content	Essay 7 = 100% human-generated content	Essay 7 = 0% AI GPT	Essay 7 = 0% AI GPT	Essay 7 = 0% AI GPT	Essay 7 = 0% AI GPT
Essay 8 = 98% human-generated content	Essay 8 = 96% human-generated content	Essay 8 = 100% human-generated content	Essay 8 = 69% human-generated content	Essay 8 = 0% AI GPT	Essay 8 = 0% AI GPT	Essay 8 = 0% AI GPT	Essay 8 = 0% AI GPT
Essay 9 = 100% human-generated content	Essay 9 = 91% human-generated content	Essay 9 = 99% human-generated content	Essay 9 = 100% human-generated content	Essay 9 = 0% AI GPT	Essay 9 = 0% AI GPT	Essay 9 = 15.37% AI GPT	Essay 9 = 3.97% AI GPT
Essay 10 = 100% human-generated content	Essay 10 = 100% human-generated content	Essay 10 = 100% human-generated content	Essay 10 = 99% human-generated content	Essay 10 = 0% AI GPT	Essay 10 = 0% AI GPT	Essay 10 = 8.74% AI GPT	Essay 10 = 0% AI GPT

In terms of false positives, Writer had two false positives and three false positives for the 2nd-year module's L1 and L2 essay sets, respectively. The first set collectively had 5% AI content, with an average false positive percentage of 2.5% AI content, while the second set contained 18% AI content, with an average false positive percentage of 6% AI content. With regard to the 3rd-year module, the L1 essay set consisted of four false positives that contained an overall AI content of 74%. Collectively, they had an average false positive percentage of 18.5% AI content. Its counterpart L2 essay set had four false positives, whose aggregate AI content was 53%. Its average false positive percentage was 10.6% AI content.

For ZeroGPT, the 2nd-year module's L1 and L2 essay sets had one false positive and no false positive, respectively. The first set contained 6.88% AI content, which was also its average false positive percentage. The second set had 0% AI content and 0% AI content as its average false positive percentage. ZeroGPT's 3rd-year module's L1 and L2 essay sets had six false positives and two false positives each. The first set had an aggregate AI content of 77.82%, with 12.97% as its average false positive percentage for its AI content. By contrast, the second essay set contained an overall AI content of 23.39%, with 11.695% being its average false positive percentage for its AI content (see Table 3).

Table 3: How the AI and human content probability scores and means were calculated.

Writer
<b>2<sup>nd</sup>-Year Module</b>
<b>L1</b> 100% human-generated content = 8 (Essays 1, 2, 3, 5, 6, 7, 9, and 10)
Human-generated content less than 100% = 2 (Essay 4 = 97% human-generated content; Essay 8 = 98% human-generated content)
False positives: 100% - 97% = 3%; 100% - 98% = 2%; + sum of two essays = 5% Average false positive percentage: 5%/2 = 2.5%. NB: 2.5%/10 = 0.25% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 2/10 = 0.2 (2/10 * 100) = 20% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (8 essays) / TP (0) + TN (8 essays) + FP (2 essays) + FN (0) = 8/10 = 0.8 (80%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 8/8 + 2 = 8/10 = 0.8 (80%)
<b>L2</b> 100% human-generated content = 7 (Essays 1, 2, 3, 4, 5, 6, and 10)
Human-generated content less than 100% = 3 (Essay 7 = 95% human-generated content; Essay 8 = 96% human-generated content; Essay 9 = 91% human-generated content)
False positives: 100% - 95% = 5%; 100% - 96% = 4%; 100% - 91% = 9%; + sum of three essays = 18% Average false positive percentage: 18%/3 = 6%. NB: 18%/10 = 1.8% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 3/10 = 0.3 (3/10 * 100) = 30% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (7 essays) / TP (0) + TN (7 essays) + FP (3 essays) + FN (0) = 7/10 = 0.7 (70%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 7/7 + 3 = 7/10 = 0.7 (70%)
<b>3<sup>rd</sup>-Year Module</b>
<b>L1</b> 100% human-generated content = 6 (Essays, 1, 3, 4, 6, 8, and 10)
Human-generated content less than 100% = 4 (Essay 2 = 98% human-generated content; Essay 5 = 94% human-generated content; Essay 7 = 35% human-generated content; Essay 9 = 99% human-generated content)
False positives: 100% - 98% = 2%; 100% - 94% = 6%; 100% - 35% = 65%; 100% - 99% = 1%; + sum of four essays = 74% Average false positive percentage: 74%/4 = 18.5%. NB: 18.5%/10 = 1.85% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 4/10 = 0.4 (4/10 * 100) = 40% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (6 essays) / TP (0) + TN (6 essays) + FP (4 essays) + FN (0) = 6/10 = 0.6 (60%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 6/6 + 4 = 6/10 = 0.6 (60%)
<b>L2</b> 100% human-generated content = 5 (Essays 1, 2, 6, 7, and 9)
Human-generated content less than 100% = 5 (Essay 3 = 81% human-generated content; Essay 4 = 99% human-generated content; Essay 5 = 99% human-generated content; Essay 8 = 69% human-generated content; Essay 10 = 99% human-generated content)
False positives: 100% - 81% = 19%; 100% - 99% = 1%; 100% - 99% = 1%; 100% - 69% = 31%; 100% - 99% = 1%; + sum of five essays = 53% Average false positive percentage: 53%/5 = 10.6%. NB: 53%/10 = 5.3% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 5/10 = 0.5 (5/10 * 100) = 50% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (5 essays) / TP (0) + TN (5 essays) + FP (5 essays) + FN (0) = 5/10 = 0.5 (50%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 5/5 + 5 = 5/10 = 0.5 (50%)
<b>ZeroGPT</b>
<b>2<sup>nd</sup>-Year Module</b>
<b>L1</b> 0% AI GPT = 9 essays Essay classified as containing miscellaneous AI content percentages = (Essay 4 = 6.88% AI GPT)
False positives: 6.88% + sum of one essay = 6.88% Average false positive percentage: 6.88%/1 = 6.88%. NB: 6.88%/10 = 0.688% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 1/10 = 0.1 (1/10 * 100) = 10% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (9 essays) / TP (0) + TN (9 essays) + FP (1 essay) + FN (0) = 9/10 = 0.9 (90%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 9/9 + 1 = 9/10 = 0.9 (90%)
<b>L2</b> 0% AI GPT = 10 essays Essay classified as containing miscellaneous AI content percentages = 0 (None)
False positives: None (0%) Average false positive percentage: None (0%) False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 0/10 = 0.0 (0/10 * 100) = 0% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (10 essays) / TP (0) + TN (10 essays) + FP (0 essay) + FN (0) = 10/10 = 1 (100%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 10/10 + 0 = 10/10 = 1 (100%)

3<sup>rd</sup>-Year Module

<b>L1</b> 0% AI GPT = 4 (Essays 4, 5, 7, and 8) Essays classified as containing miscellaneous AI content percentages = 6 (Essay 1 = 14.91% AI GPT; Essay 2 = 5.35% AI GPT; Essay 3 = 16.49% AI GPT; Essay 6 = 16.96% AI GPT; Essay 9 = 15.37% AI GPT; Essay 10 = 8.74% AI GPT)
False positives: 14.91%; 5.35%; 16.49%; 16.96%; 15.37%; 8.74% + sum of six essays = 77.82% Average false positive percentage: 77.82%/6 = 12.97%. NB: 12.97%/10 = 1.297% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 6/10 = 0.6 (6/10 * 100) = 60% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (4 essays) / TP (0) + TN (4 essays) + FP (6 essays) + FN (0) = 4/10 = 0.4 (40%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 4/6 + 4 = 4/10 = 0.4 (40%)
<b>L2</b> 0% AI GPT = 8 essays Essays classified as containing miscellaneous AI content percentages = 2 (Essay 1 = 19.42% AI GPT; Essay 9 = 3.97% AI GPT)
False positives: 19.42%; 3.97% = 23.39% + sum of two essays = 23.39% Average false positive percentage: 23.39%/2 = 11.695%. NB: 23.39%/10 = 2.339% False positive rate (FPR) = incorrectly detected AI samples (essays)/all human-written samples (essays) = 2/10 = 0.2 (2/10 * 100) = 20% Accuracy OR true negative rate: correctly detected essays/all essays, or, TP (0) + TN (8 essays) / TP (0) + TN (8 essays) + FP (2 essays) + FN (0) = 8/10 = 0.8 (80%) True negative rate (TNR) = correctly detected human-written samples (essays)/all human-written samples (essays) = 8/8 + 2 = 8/10 = 0.8 (80%)

Since the raw false positives and their corresponding average false positive percentages as discussed above are not a reliable measure of the accuracy of AI detectors, false positive rates (FPRs), true negative rates (TNRs), and the accuracy of the scan scores of the 30 AI detectors for the four sets of essays were calculated (see Captain Words, 2024; Colquhoun, 2014; Gillham, 2024; Weber-Wulff et al., 2023; Wu et al., 2023; also see Table 3). In particular, the FPRs, the TNRs, the accuracy, and the specificity of the AI detectors whose scan scores were direct opposites of each other, were chosen and calculated for comparative analysis. Included in the 30 AI detectors are the AI detectors that correctly classified all ten essays in each of the four essay sets (two sets for English L1 and two sets for English L2), which were tested by the 30 AI detectors. They also encompassed the AI detectors that completely misclassified all ten essays in each of these four essay sets. In this context, two AI detectors, Copyleaks and Undetectable AI, correctly classified all ten essays in each of the four essay sets (see Table 4). Contrariwise, nine AI detectors completely misclassified all ten essays in each of these four essay sets. These nine AI detectors were AI Content Checker, AI-Detector, AI Detector, Detecting-AI.com, GLTR, GPT-2 Output Detector Demo, IvyPanda GPT Essay Checker, RewriteGuru's AI Detector, and SEO (see Table 5).

Table 4: How Copyleaks and Undetectable AI correctly detected all the essay sets in both English language categories of the two modules.

Copyleaks				Undetectable AI			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n = 10)	L2 (n = 10)	L1 (n = 10)	L2 (n = 10)	L1 (n = 10)	L2 (n = 10)	L1 (n = 10)	L2 (n = 10)
FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)
Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)
TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)
AI Content Detector				Hive Moderation			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n = 10)	L2 (n = 10)	L1 (n = 10)	L2 (n = 10)	L1 (n = 10)	L2 (n = 10)	L1 (n = 10)	L2 (n = 10)
FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0 (0%)	FPR = 0.3 (30%)	FPR = 0.1 (10%)	FPR = 0 (0%)	FPR = 0.1 (10%)	FPR = 0 (0%)
Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 1 (100%)	Accuracy = 0.7 (70%)	Accuracy = 0.9 (90%)	Accuracy = 1 (100%)	Accuracy = 0.9 (90%)	Accuracy = 1 (100%)
TNR = 1 (100%)	TNR = 1 (100%)	TNR = 1 (100%)	TNR = 0.7 (70%)	TNR = 0.9 (90%)	TNR = 1 (100%)	TNR = 0.9 (90%)	TNR = 1 (100%)

The three measures: the FPR (false positive rate), accuracy, and the TNR (true negative rate) were manually calculated based on the scan scores of the said AI detectors. The FPR was calculated using the formula, FPR = incorrectly detected AI essays/all human-written essays, or FP/FP + TP, where

FP and TP stand for false positives and true positives, respectively. This is related to each essay set (see Table 3). In the same breadth, accuracy was calculated by utilising the formula, accuracy = correctly detected essays/all essays, or  $TP + TN / TP + TN + FP + FN$ . In this case, TN and FN stand for true negatives and false negatives. For its part, the TNR was calculated through this formula:  $TNR = \text{correctly detected human-written essays} / \text{all human-written essays}$ , or  $TN / TN + FP$  (see Table 3). For example, Table 4 depicts the FPR, the accuracy, and the TNR of each of the L1 and L2 essay sets of both the 2nd-year module and the 3rd-year module for Writer and ZeroGPT. On one hand, for the 2nd-year module's L1 and L2, Writer had the following sets of scores for each of these two English language categories: FRP = 0.2, Accuracy = 0.8, and TNR = 0.8; and FRP = 0.3, Accuracy = 0.7, and TNR = 0.7. Its 3rd-year module's L1 and L2 scores for these three measures were as follows: FRP = 0.4, Accuracy = 0.6, and TNR = 0.6; and FRP = 0.5, Accuracy = 0.5, and TNR = 0.5.

Table 5: How the nine AI incorrectly detected all the essay sets in both English language categories of the two modules.

AI-Detector				AI Detector			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)
FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)
Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)
TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)

Detecting-AI.com				GLIX			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)
FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)
Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)
TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)

GPT-2 Output Detector Demo				IvyPanda GPT Essay Checker			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)
FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)
Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)
TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)

RewriteGuru's AI Detector				SEO.AI			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module		2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)
FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)
Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)
TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)

AI Content Checker			
2 <sup>nd</sup> -Year Module		3 <sup>rd</sup> -Year Module	
L1 (n=10)	L2 (n=10)	L1 (n=10)	L2 (n=10)
FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)	FPR = 1 (100%)
Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)	Accuracy = 0 (0%)
TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)	TNR = 0 (0%)

On the other hand, ZeroGPT had the following score sets for its 2nd-year module's L1 and L2: FRP = 0.1, Accuracy = 0.9, and TNR = 0.9; and FRP = 0.0, Accuracy = 1, and TNR = 1. And its score sets for the 3rd-year module's L1 and L2 were as follows: FRP = 0.6, Accuracy = 0.4 and TNR = 0.4; and FRP = 0.2, Accuracy = 0.8, and TNR = 0.8. With the exception of two essay sets (the 2nd-year module's L1 for Writer and the 3rd-year module's L2 for ZeroGPT), the two AI detectors had varying scores for these three measures in their other essay sets for these two modules. Suffice it to say that ZeroGPT correctly classified one essay set for the 2nd-year module's L2, while it incorrectly identified this module's L1 by one percentage point. Therefore, ZeroGPT performed

better between the two AI detectors.

The points discussed in the preceding paragraph, lead to the calculation of the FPRs, the TNRs, the accuracy, and the specificity of the two AI detectors that correctly identified all the essay sets and of the nine AI detectors that incorrectly identified all the essay sets. Specificity is the function of TNR: it is about the proportion of correct/true negative cases correctly classified as such by an AI detector (see Elkhata et al., 2023). In the context of the present study, this relates to the proportion of student-written essays correctly recognised by any of the 30 AI detectors out of ten student-written essays in each of the four essay sets. To calculate these four measures in the two sets of AI detectors mentioned above, an online confusion matrix calculator was used. This calculator was ideal for computing these measures. As said earlier, for Copyleaks, Undetectable AI, and the other nine AI detectors, the scores are as portrayed in Table 6.

Table 6: FPRs, TNRs, the accuracy, and the specificity of Copyleaks and Undetectable AI (top half) and of the other nine AI detectors (bottom half) for English L1 and English L2 essay sets as measured by a confusion matrix calculator.

Measure	Value	Formula
Sensitivity	NAN	$TPR = TP / (TP + FN)$
Specificity	1	$SPC = TN / (FP + TN)$
Negative Predictive Value	1	$NPV = TN / (TN + FN)$
False Positive Rate	0	$FPR = FP / (FP + TN)$
False Negative Rate	NAN	$FNR = FN / (FN + TP)$
Accuracy	1	$ACC = (TP + TN) / (TP + TN + FP + FN)$

Measure	Value	Formula
Sensitivity	NAN	$TPR = TP / (TP + FN)$
Specificity	0	$SPC = TN / (FP + TN)$
Negative Predictive Value	NAN	$NPV = TN / (TN + FN)$
False Positive Rate	1	$FPR = FP / (FP + TN)$
False Negative Rate	NAN	$FNR = FN / (FN + TP)$
Accuracy	0	$ACC = (TP + TN) / (TP + TN + FP + FN)$

As depicted in the top half of this table, the scores for the FPR, the negative predictive value (NPV) (which is also an equivalent of a true negative rate (TNR)), accuracy, and specificity for both Copyleaks and Undetectable AI were as follows: FPR = 0, NPV = 1, accuracy = 1, and specificity = 1. The acronym, NAN (not a number), or sometimes, NaN, denotes the measures whose scores could not be computed as they were not relevant for the purpose at hand. As was highlighted concerning Table 4 earlier, Copyleaks and Undetectable AI had these scores because they correctly identified all of the essay sets which consisted of the two English language categories. Inversely, as exhibited in the bottom half of Table 6, the nine AI detectors mentioned above, collectively had the score set, FPR = 1, accuracy = 0, and specificity = 0, since all of them misidentified all the essay sets of the two English language categories for both modules. Here, too, NAN signifies the measures whose scores could not be captured as they were not relevant.

All the 30 AI detectors were ranked for their accuracy in detecting if the four sets of essays (two sets of English L1 essays, n = 20; and two sets of English L2 essays, n = 20) were GenAI-generated or human-written. The accuracy and TNR

scores of each AI detector were used to rank the accuracy of the 30 AI detectors (for relevant examples, see Tables 3 and 4). Based on these composite scores, many AI detectors shared joint spots when they were ranked for accuracy. For instance, two AI detectors, Copyleaks and Undetectable AI, jointly shared the first spot. They were followed by Hive Moderation and Scribbr, AI Content Detector and Plagiarism Detector, and Dupli Checker and Grammarly, which, as pairs, jointly shared the second, third, and fourth spots, respectively. ZeroGPT and Detect Bard, each notched the fifth and sixth places, while AI Checker Tool and AI Contentfy jointly occupied the seventh position. This is followed by Writer in the eighth spot and Rank Wizard AI and Sapling jointly took up the ninth position.

The spots ranging from ten to 15 were, each, occupied by different AI detectors, with GPTZero at the tenth spot and QuillBot AI Detector at the 15th place. The 16th and last spot was collectively shared by the nine AI detectors mentioned earlier.

## Discussion

The results presented above is discussed in this section in response to the three research questions for this study.

### The accuracy of 30 AI content detectors

As highlighted in the preceding section, of the 30 free, publicly available online AI detectors, only two of them, Copyleaks and Undetectable AI, were able to correctly identify all the essay sets of the two English language categories (English L1 and English L2) as human written. These two AI detectors also had the highest accuracy and TNR scores for all these essay sets, when their scores were manually calculated. Moreover, they did so even when their specificity and NPV was computed using a confusion matrix calculator. However, their scores in all these four measures diametrically contrasted with those of the nine AI detectors, whose scores in these measures, especially for accuracy and specificity, were zero (0%). Their FPR score of one (100%) was the polar opposite of the FPR score of zero (0%) for Copyleaks and Undetectable AI. In this sense, the nine AI detectors misidentified all four essay sets of the two English language categories. The rest of the other AI detectors had varying accuracy, FPR, and TNR scores. As such, they classified these four essay sets of English L1 and English L2 in varying degrees of accuracy, FPRs, and TNRs (see Figure 1).

In some of the previous studies conducted on the efficacy of AI detectors, Copyleaks has been the best-performing AI detector or, at least one of the best-performing AI detectors. One such study is Walters' (2023) study. This study tested the effectiveness of 16 AI detectors in identifying GenAI-generated and human-written content in three sets of first-year, undergraduate composition essays. The three sets of essays comprised 42 essays generated by ChatGPT-3.5, 42 essays created by ChatGPT-4, and 42 essays written by students. The last set was chosen from a college's English 110 (First-Year Composition) essays, which had been

submitted during the 2014-2015 academic year. In this study, both Copyleaks and Turnitin had the highest accuracy rate, followed by Originality. Sapling and Content at Scale had the lowest accuracy rate among the 16 AI detectors. In the current study, Sapling and Content at Scale, occupied the 9th and 13th spots respectively.

Another study is Chaka's (2023a), which evaluated the accuracy of five AI detectors in detecting GenAI-generated content in 21 applied English language studies responses generated by three GenAI chabots: ChatGPT (n = 6), YouChat (n = 7), and Chatsonic (n = 8). The five AI detectors were GPTZero, OpenAI Text Classifier, Writer, Copyleaks, and GLTR. All the twenty-one English responses were submitted to the five AI detectors for scanning. The ChatGPT-generated responses were translated into German, French, Spanish, Southern Sotho, and isiZulu by using Google Translate. They were, then, submitted to GPTZero for scanning. The German, French and Spanish translated versions were inputted into Copyleaks for scanning. In this sense, this study utilised machine translation as an adversarial attack, which is a strategy that is related to a critical studies approach to AI as I had argued in the relevant section above. In all the different versions of the twenty-one responses, Copyleaks was the most accurate of the five AI detectors (see Chaka, 2023a). Similarly, in a literature and integrative hybrid review conducted by Chaka (2024), which reviewed 17 peer-reviewed journal articles, Copyleaks was one of the best-performing AI detectors in one of the four articles in which OpenAI Text Classifier, and Crossplag, Grammarly also topped in each of the other three articles. But, overall, in all the 17 reviewed articles, Crossplag was the best-performing AI detector, followed by Copyleaks.

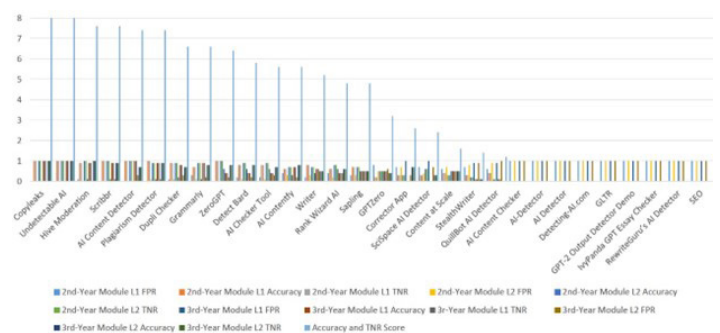


Figure 1: A graphic representation of the 30 AI detectors based on their accuracy, FPR, and TNR scores.

In Odri and Yoon's (2023) study, though, which as discussed earlier, tested 11 AI detectors and employed adversarial attacks, especially evasion techniques (e.g., improving command messages (prompts) in ChatGPT, adding minor grammatical errors, paraphrasing, and substituting Latin letters with their Cyrillic equivalents), as part of a critical studies approach to AI, Originality out-performed all the 11 AI detectors in correctly identifying the human-written-text. It, nonetheless, misclassified the final version of the AI-generated text. However, it was the AI detector that was most resistant to adversarial attacks compared to the other AI detectors (see Odri & Yoon, 2023).

## **Differential classification of the four sets of student essays and a language category assigned more false positives**

As pointed out in the preceding section, both CopyLeaks and Undetectable AI classified all the four sets of English L1 and English L2 student essays similarly and correctly by assigning the same scores for the three measures: accuracy, FPR, and TNR, to all of them. Additionally, both did so for their specificity scores for all the four essay sets. Likewise, the nine AI detectors allotted the same scores for their respective measures for the four essay sets. Even the rest of the other AI detectors, which had varying scores for these measures, did not have scores specifically skewed toward one English language category in each of the four essay sets. In fact, even in the cases where one AI detector had lower scores for essays within a given essay set of a particular language category, it had higher scores for essays within another essay set of a different language category.

In instances where a particular AI detector scored the essay sets of the one language category in a given module (e.g., the English L1 essay sets for both the 2nd-year module and the 3rd-year module) higher than the essay sets of the other language category in the same module, the differences in the scores of essay sets of these different language categories were not substantial. Or, if the scores were higher, they were not consistent for the essay sets of one language category (e.g., English L1) to the exclusion of the essay sets of the other language category (e.g., English L2) (see Table 3). So, in the present study, the AI detectors that correctly classified the student essay sets did so for both English L1 and English L2. In a similar vein, those AI detectors that misclassified the student essay sets did so for both of these English language categories. Moreover, no language category was assigned more false positives for its essay sets than those of the other language category. This means that the 30 AI detectors were not language category-biased or language category-sensitive when assigning false positives to and when classifying the essays belonging to the four essay sets. In the current study, therefore, there is no evidence suggesting that the AI detectors that were tested were consistently and invariably biased towards or against any of the student essay sets of the two English language categories.

In contrast, though, and as stated earlier, Liang et al.'s (2023) study found that the AI detectors that they evaluated tended to be biased against non-English language speakers' essays (also see Mathewson, 2023; Shane, 2023; cf. Adamson, 2023; Gillham, 2024). While this is the case, the results of the current study, nonetheless, do not nullify or invalidate those of Liang et al.'s (2023) study, as it did not use the same data sets as the ones used by that study. Instead, the present study's results are different from those of Liang et al.'s (2023) study.

## **Implications and recommendations**

This study has implications for detecting GenAI content in student essays and for differentiating between GenAI-generated and human-written content in student essays in higher education. Firstly, detecting GenAI in student essays

or distinguishing between GenAI-generated and human-written content in such essays is not simply a matter of displaying AI and human content probability scores (or percentages) and the statements accompanying them as most, if not all, AI content detectors currently tend to do. Neither is it a matter of making self-serving claims about high AI detection accuracy rates, as is the case with 28 (93%) of the 30 AI detectors tested in this study. This means that the AI detection accuracy claims made by different AI detection tools on their respective landing pages should be taken with a pinch of salt. As demonstrated in this study, such claims hardly live up to their stated expectations. Again, as shown by the results of this study, of these 28 AI detectors that did not perform as expected, nine of them completely misclassified all the human-written essays, while the remaining 19 misclassified these essays in varying degrees. Any AI content probability percentage or percentage point, however negligible it may be, that is attributed to a student essay which has no GenAI content at all, inflicts immeasurable reputational damage to that essay and to the student who produced it. This means that if this particular essay was meant for assessment purposes, then, the student concerned would be unfairly accused of a gross academic dishonesty they would not have committed. Given all of this, it is advisable for academics and for universities to which these academics belong, to exercise extreme caution when utilising any AI content detection tool for detecting GenAI content in their students' academic essays. The reason for having to be extra cautious is that most of the current AI detectors demonstrate a high degree of inaccuracy and unreliability. Importantly, it is very risky to employ one AI content detector and take its scan results as a final verdict for any given human-written text.

Secondly, the reliance of the current AI detectors on perplexity and burstiness for determining and predicting the presence or absence of GenAI content in human-written student essays results in these detectors consistently misclassifying such essays. This is one of the reasons why they keep on misclassifying student writing that has low perplexity and burstiness, such as that of non-English native speakers, as containing GenAI content portions, even when that is not the case. Repetitive word sequences and predictable lexical and syntactic parsing, as assumed by perplexity and burstiness, might work as indicators of the presence or the absence of GenAI content within the surreal world of GenAI driven by large language models. Nevertheless, in a real-world and human environment in which university students produce different forms of academic writing, informed by their diverse English language backgrounds and in response to assignment questions, perplexity and burstiness serve as weak, if not misplaced, indicators of the presence or the absence of GenAI content in student writing. The types of essays used in the current study serve as a case in point that detecting GenAI-generated content or distinguishing between it and its human-written counterpart is not merely a matter of English L1 writing versus English L2 writing. Human-produced writing cannot be reduced to robotic writing powered and aided by machine learning and GenAI large language models. Therefore, it is prudent for AI detection tools to have language training data sets that reflect the diverse, multi-dialectal, poly-racial, and pluri-ethnic speakers of a given language, in various global or



geographical settings, for them to be able to capture the nuances of such a language. This is more so for a language such as English that has these types of speakers across the globe.

## Conclusion

The current study had three research questions (RQs) and three corresponding objectives as stated earlier. Only two of the 30 tested, free-to-use, AI detectors, Copyleaks and Undetectable AI, did manage to correctly detect all of the student essay sets of the two English language categories (English L1 and English L2) as human-written. Nine of these 30 AI detectors (AI Content Checker, AI-Detector, AI Detector, Detecting-AI.com, GLTR, GPT-2 Output Detector Demo, IvyPanda GPT Essay Checker, RewriteGuru's AI Detector, and SEO) did the opposite: they misidentified all the essays in each of the four essay sets of the two language categories in both the 2nd-year module and the 3rd-year module. The remaining 19 AI detectors both correctly and incorrectly classified the four essay sets in varying degrees without any bias to any essay set of the two English language categories. Therefore, Copyleaks and Undetectable AI, were, jointly, the top-most accurate AI detectors that ranked first in this study, while the nine AI detectors were the most inaccurate, which collectively ranked last in the pecking order. Of the other 19 AI detectors, ten of them held joint positions, with the remaining nine notching individual accuracy slots in the ranking.

All 30 AI detectors did not assign differential classification to the four essay sets according to the English language categories to which they belonged. That is, they displayed no specific bias toward language categories in classifying or misclassifying the four essay sets. The same applies to the false positives they accorded to these essay sets. If only two AI detectors out of 30 can accurately detect all the student essay sets across the two language categories, and nine AI detectors can do the complete opposite, with the remaining AI detectors yielding variable accuracy scores for the same sets of essays in the two language categories as is the case in this study, then, university students and the universities to which they belong are in trouble concerning the presence or absence of GenAI content in student essays. Moreover, the results of the current study demonstrate that detecting GenAI-generated content or distinguishing it from its human-written counterpart is not simply a matter of perplexity and burstiness, or of English L1 writing versus English L2 writing. Human-produced writing is very complex and nuanced and thus cannot be reduced to measures of high or low perplexity and burstiness. This applies to both English L1 and English L2 writers, depending on their writing proficiency. On this basis, the present study suggests that the bulk of the currently available AI detectors are not fit for its purpose, even when the input content, such as the essays used in this study, is not manipulated through any adversarial attacks. The implication of this study, therefore, is that relying on one or even a few AI detection tools for identifying GenAI content in student essays is a risky move.

## References

- Abani, S., Volk, H. A., De Decker, S., Fenn, J., Rusbridge, C., Charalambous, M., ... Nessler J. N. (2023). ChatGPT and scientific papers in veterinary neurology: Is the genie out of the bottle? *Frontiers in Veterinary Science*, 10(1272755), 1-7. <https://doi.org/10.3389/fvets.2023.1272755>
- Adamson, D. (2023). *New research: Turnitin's AI detector shows no statistically significant bias against English language learners*. <https://www.turnitin.com/blog/new-research-turnitin-s-ai-detector-shows-no-statistically-significant-bias-against-english-language-learners>
- AIContentfy Team. (2023). *Evaluating the effectiveness of AI detectors: Case studies and metrics*. <https://aicontentfy.com/en/blog/evaluating-of-ai-detectors-case-studies-and-metrics>
- Alexander, K., Savvidou, C., & Alexander, C. (2023). Who wrote this essay? Detecting AI-generated writing in second language education in higher education. *The Journal of Teaching English with Technology*, 23(20), 25-43. <https://doi.org/10.56297/BUKA4060/XHLD5365>
- Anderson, C. (2023). *The false promise of AI writing detectors*. <https://www.linkedin.com/pulse/false-promise-ai-writing-detectors-carol-anderson>
- Anil, A., Saravanan, A., Singh, S., Shamim, M. A., Tiwari, K., Lal, H., ...Sah, R. (2023). Are paid tools worth the cost? A prospective cross-over study to find the right tool for plagiarism detection. *Heliyon*, 9(9), e19194, 1-11. <https://doi.org/10.1016/j.heliyon.2023.e19194>
- Anthology White Paper. (2023). *AI, academic integrity, and authentic assessment: An ethical path forward for education*. [https://www.anthology.com/sites/default/files/2023-09/White%20Paper-AI%20Academic%20Integrity%20and%20Authentic%20Assessment-An%20Ethical%20Path%20Forward%20for%20Education-v2\\_09-23\\_0.pdf](https://www.anthology.com/sites/default/files/2023-09/White%20Paper-AI%20Academic%20Integrity%20and%20Authentic%20Assessment-An%20Ethical%20Path%20Forward%20for%20Education-v2_09-23_0.pdf)
- Bisi, T., Risser, A., Clavert, P., Migaud, H., & Dartus, J. (2023). What is the rate of text generated by artificial intelligence over a year of publication in orthopedics and traumatology: Surgery and research? Analysis of 425 articles before versus after the launch of ChatGPT in November 2022. *Orthopaedics and Traumatology: Surgery and Research*, 109(8), 103694. <https://doi.org/10.1016/j.otsr.2023.103694>
- Blau, I., Goldberg, S., Friedman, A., & Eshet-Alkalai, Y. (2020). Violation of digital and analog academic integrity through the eyes of faculty members and students: Do institutional role and technology change ethical perspectives? *Journal of Computing in Higher Education*, 33(1), 157-187. <https://doi.org/10.1007/s12528-020-09260-0>
- Cai, S., & Cui, W. (2023). *Evade ChatGPT detectors via a single space*. <https://arxiv.org/pdf/2307.02599.pdf>
- Captain Words. (2024). *Testing AI detection tools – our methodology*. <https://captainwords.com/ai-detection-tools-test-methodology/>

- Chaka, C. (2022). Digital marginalization, data marginalization, and algorithmic exclusions: A critical southern decolonial approach to datafication, algorithms, and digital citizenship from the Souths. *Journal of e-Learning and Knowledge Society*, 18(3), 83-95. <https://doi.org/10.20368/1971-8829/1135678>
- Chaka, C. (2023a). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning & Teaching*, 6(2), 94-104. <https://doi.org/10.37074/jalt.2023.6.2.12>
- Chaka, C. (2023b). Generative AI chatbots - ChatGPT versus YouChat versus Chatsonic: Use cases of selected areas of applied English language studies. *International Journal of Learning, Teaching and Educational Research*, 22(6), 1-19. <https://doi.org/10.26803/ijlter.22.6.1>
- Chaka, C. (2024). Reviewing the performance of AI detection tools in differentiating between AI-generated and human-written texts: A literature and integrative hybrid review. *Journal of Applied Learning & Teaching*, 7(1), 1-12. <https://doi.org/10.37074/jalt.2024.7.1.14>
- Cingillioglu, I. (2023). Detecting AI-generated essays: The ChatGPT challenge. *The International Journal of Information and Learning Technology*, 40(3), 259-268. <https://doi.org/10.1108/IJILT-03-2023-0043>
- Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*, 1(140216), 1-16. <https://doi.org/10.1098/rsos.140216>
- Couldry, N., & Mejias, U. A. (2019a). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media*, 20(4), 336-349. <https://doi.org/10.1177/1527476418796632>
- Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport*, 40(2), 615-622. <https://doi.org/10.5114/biolSport.2023.125623>
- Desaire, H. A., Chua, A. E., Isom, M., Jarosova, R., & Hua, D. (2023). Distinguishing academic science writing from humans or ChatGPT with over 99% accuracy using off-the-shelf machine learning tools. *Cell Reports Physical Science*, 4(6), 1-2. <https://doi.org/10.1016/j.xcrp.2023.101426>
- Elali, F. R., & Rachid, L. N. (2023). AI-generated research paper fabrication and plagiarism in the scientific community. *Patterns*, 4, 1-4. <https://doi.org/10.1016/j.patter.2023.100706>
- Elkhatat, A. M., Elsaid, K., & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI generated text. *International Journal for Educational Integrity*, 19(17), 1-16. <https://doi.org/10.1007/s40979-023-00140-5>
- Ferrara, E. (2023). *Should ChatGPT be biased? Challenges and risks of bias in large language models*. <https://arxiv.org/abs/2304.03738>
- Gamage, K. A. A., De Silva, E. K., & Gunawardhana, N. (2020). Online delivery and assessment during COVID-19: Safeguarding academic integrity. *Education Sciences*, 10(301), 1-24. <https://doi.org/10.3390/educsci10110301>
- Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, 6(75), 1-5. <https://doi.org/10.1038/s41746-023-00819-6>
- Gillham, J. (2024). *Native English speakers?* <https://originality.ai/blog/are-gpt-detectors-biased-against-non-native-english-speakers>
- Habibzadeh, F. (2023). GPTZero performance in identifying artificial intelligence-generated medical texts: A preliminary study. *Journal of Korean Medical Sciences*, 38(38), e319. <https://doi.org/10.3346/jkms.2023.38.e319>
- Homolak, J. (2023). Exploring the adoption of ChatGPT in academic publishing: Insights and lessons for scientific writing. *Croatian Medical Journal*, 64(3), 205-207. <https://doi.org/10.3325/cmj.2023.64.205>
- Ibrahim, K. (2023). Using AI based detectors to control AI assisted plagiarism in ESL writing: "The terminator versus the machines". *Language Testing in Asia*, 13(46), 1-28. <https://doi.org/10.1186/s40468-023-00260-2>
- Ifelebuegu, A. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning and Teaching*, 6(2), 385-392. <https://doi.org/10.37074/jalt.2023.6.2.2>
- Krishna, K., Song, Y., Karpinska, M., Wieting, J., & Iyer, M. (2023). *Paraphrasing evades detectors of AI-generated text, but retrieval is an effective defense*. <https://arxiv.org/abs/2303.13408>
- Ladha, N., Yadav, K., & Rathore, P. (2023). AI-generated content detectors: Boon or bane for scientific writing. *Indian Journal of Science and Technology*, 16(39), 3435-3439. <https://doi.org/10.17485/IJST/v16i39.1632>
- Lee, D. (2023). *How hard can it be? Testing the reliability of AI detection tools*. [https://www.researchgate.net/profile/Daniel-Lee-95/publication/374170650\\_How\\_hard\\_can\\_it\\_be\\_Testing\\_the\\_reliability\\_of\\_AI\\_detection\\_tools/links/6512b65237d0df2448edc358/How-hard-can-it-be-Testing-the-reliability-of-AI-detection-tools.pdf](https://www.researchgate.net/profile/Daniel-Lee-95/publication/374170650_How_hard_can_it_be_Testing_the_reliability_of_AI_detection_tools/links/6512b65237d0df2448edc358/How-hard-can-it-be-Testing-the-reliability-of-AI-detection-tools.pdf)
- Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. *Patterns*, 4(7), 1-4. <https://doi.org/10.1016/j.patter.2023.100779>
- Lindgren, S. (2023). Introducing critical studies of artificial intelligence. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 1-19). Cheltenham: Edward Elgar Publishing. <http://dx.doi.org/10.4337/9781803928562.00005>

- Makri, C., & Neely, A. (2021). Grounded theory: A guide for exploratory studies in management research. *International Journal of Qualitative Methods*, 20, 1-14. <https://doi.org/10.1177/16094069211013654>
- Martin, J. L., & Wright, K. E. (2023). Bias in automatic speech recognition: The case of African American language. *Applied Linguistics*, 44(4), 613-630. <https://doi.org/10.1093/applin/amac066>
- Mathewson, T. G. (2023). *AI detection tools falsely accuse international students of cheating*. The Markup. <https://themarkup.org/machine-learning/2023/08/14/ai-detection-tools-falsely-accuse-international-students-of-cheating>
- Mohamed, S., Png, M.-T., & Isaac, W. (2020). Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology*, 33, 659-684. <https://doi.org/10.1007/s13347-020-00405-8>
- Odri, G. A., & Yoon, D. J. Y. (2023). Detecting generative artificial intelligence in scientific articles: Evasion techniques and implications for scientific integrity. *Orthopaedics & Traumatology: Surgery & Research*, 109(8), 103706. <https://doi.org/10.1016/j.otsr.2023.103706>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice*, 20(2). <https://doi.org/10.53761/1.20.02.07>
- Perkins, M., Roe, J., Postma, D., McGaughan, J., & Hickerson, D. (2024). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics*, 22, 89-113. <https://doi.org/10.1007/s10805-023-09492-6>
- Ricaurte, P. (2019). Data epistemologies, the coloniality of power, and resistance. *Television & New Media*, 20(4), 350-365. <https://doi.org/10.1177/1527476419831640>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sadasivan, V. S., Kumar, A., Balasubramanian, S., Wang, W., & Feizi, S. (2023). *Can AI-generated text be reliably detected?* <https://arxiv.org/abs/2303.11156>
- Santra, P. P., & Majhi, D. (2023). Scholarly communication and machine-generated text: Is it finally AI vs AI in plagiarism detection? *Journal of Information and Knowledge*, 60(3), 175-183. <https://doi.org/10.17821/srels/2023/v60i3/171028>
- Shane, J. (2023). *Don't use AI detectors for anything important*. Fortune. <https://www.aiweirdness.com/dont-use-ai-detectors-for-anything-important/>
- Sobaih, A. E. E. (2024). Ethical concerns for using artificial intelligence chatbots in research and publication: Evidences from Saudi Arabia. *Journal of Applied Learning & Teaching*, 7(1), 1-11. <http://journals.sfu.ca/jalt/index.php/jalt/index>
- Subramaniam, R. (2023). Identifying text classification failures in multilingual AI-generated content. *International Journal of Artificial Intelligence and Applications (IJAlA)*, 14(5), 57-63. <https://doi.org/10.5121/ijaia.2023.14505>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 31-40. <https://journals.sfu.ca/jalt/index.php/jalt/article/view/731>
- Swedberg, R. (2020). Exploratory research. In C. Elman, J. Gerring, & J. Mahoney (Eds.), *The production of knowledge: Enhancing progress in social science* (pp. 17- 41). Cambridge: Cambridge University Press.
- Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education & Technology*, 3(1), 45-54. [https://www.researchgate.net/publication/370299956\\_ChatGPT\\_and\\_Academic\\_Integrity\\_Concerns\\_Detecting\\_Artificial\\_Intelligence\\_Generated\\_Content](https://www.researchgate.net/publication/370299956_ChatGPT_and_Academic_Integrity_Concerns_Detecting_Artificial_Intelligence_Generated_Content)
- Walters, W. H. (2023). The effectiveness of software designed to detect AI-generated writing: A comparison of 16 AI text detectors. *Open Information Science*, 7(20220158), 1-24. <https://doi.org/10.1515/opis-2022-0158>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., ... Waddington, L. (2023). *Testing of detection tools for AI-generated text*. <https://doi.org/10.48550/arxiv.2306.15666>
- Wee, H. B., & Reimer, J. D. (2023). Non-English academics face inequality via AI-generated essays and countermeasure tools. *BioScience*, 73, 476-478. <https://doi.org/10.1093/biosci/biad034>
- Wiggers, K. (2023). *Most sites claiming to catch AI-written text fail spectacularly*. TechCrunch. <https://techcrunch.com/2023/02/16/most-sites-claiming-to-catch-ai-written-text-fail-spectacularly/>
- Wu, J., Yang, S., Zhan, R., Yuan, Y., Wong, D. F., & Chao, L. S. (2023). *A survey on LLM-generated text detection: Necessity, methods, and future directions*. <https://arxiv.org/pdf/2310.14724.pdf>

Copyright: © 2024. Chaka Chaka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



---

**Artificial intelligence in higher education database (AIHE V1): Introducing an open-access repository**

---

Mohamed Fadhil bin Mohamed Ismail <sup>A</sup>	A	<i>Kaplan Higher Education Academy, Singapore</i>
Joseph Crawford <sup>B</sup>	B	<i>University of Tasmania, Australia</i>
Shannon Tan <sup>C</sup>	C	<i>Kaplan Higher Education Academy, Singapore</i>
Jürgen Rudolph <sup>D</sup>	D	<i>Kaplan Higher Education Academy, Singapore</i>
Eunice Tan <sup>E</sup>	E	<i>Singapore Institute of Management, Singapore</i>
Pauline Seah <sup>F</sup>	F	<i>Kaplan Higher Education Academy, Singapore</i>
Fiona Xiaofei Tang <sup>G</sup>	G	<i>Kaplan Business School, Perth, Australia</i>
Florence Ng <sup>H</sup>	H	<i>University Partnership (Coventry University), PSB Academy, Singapore</i>
Laura Visser Kaldenbach <sup>I</sup>	I	<i>Singapore Institute of Management, Singapore</i>
Anand Naidu <sup>J</sup>	J	<i>Murdoch University, Perth, Australia</i>
Vanessa Stafford <sup>K</sup>	K	<i>Kaplan Business School, Sydney, Australia</i>
Michelle Kane <sup>L</sup>	L	<i>University of Portsmouth, United Kingdom</i>

---

**Keywords**

AI;  
AIEd;  
Artificial intelligence;  
artificial intelligence in education;  
ChatGPT;  
GenAI;  
generative AI;  
higher education;  
literature survey;  
PRISMA.

**Abstract**

Generative artificial intelligence (GenAI) has fired the world's imagination. The higher education sector is not immune from the GenAI hype, panic, and mania. The emergence of artificial intelligence, in its newest form, into curriculum, student life, and learning has created an entanglement of technology, people, and learning. Yet, there is still a lack of cohesive accounts of the emergent literature used to inform practical learning and teaching decisions. Our manuscript responds with the deployment of a previously published systematic literature review to create the first version of the Artificial Intelligence in Higher Education Database (AIHE V1). Published in conjunction with this article, we pioneer an open-access resource to support learning and teaching scholars to gain timely access to pre-examined literature on AI and higher education. This first version documents 160 manuscripts published between 30 November 2022 and 31 December 2023. Using a rigorous systematic review method, engaging in the PRISMA approach, we offer a first glance at the metadata of articles published on AI and higher education during the first year of ChatGPT.

## Correspondence

fadhil.ismail@kaplan.com <sup>A</sup>

---

## Article Info

Received 9 March 2024

Received in revised form 28 March 2024

Accepted 3 April 2024

Available online 18 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.35>

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.35D>

## Introduction

ChatGPT's launch in late 2022 unleashed an avalanche of scholarly investigations that examine the intersections of ChatGPT, generative AI, and higher education (Rudolph et al., 2023a, 2023b). These inquiries have been disseminated across an array of academic journals and preprint repositories. Despite the high volume of these publications, they offer only fragmented views of a domain evolving at breakneck speed. Considering the rapid proliferation of such scholarly work, it is imperative to critically evaluate the corpus of existing literature. This endeavour is not merely academic. Our findings and database provide a foundation for elucidating the roles and repercussions of AI technologies within higher education contexts. Specifically, they are instrumental in identifying both the prospects and perils AI presents to teaching and learning in tertiary education (Rudolph et al., 2024).

Many authors optimistically underscore the potential of ChatGPT and similar generative AI-driven chatbots to enrich and augment educational outcomes and experiences in higher education (e.g., Rasul et al., 2023). However, there is a need to investigate GenAI's pitfalls, safeguard against unethical or ineffectual deployment, and promote its ethical, effective, and responsible use. As the body of literature expands, the importance of not only aggregating and scrutinising these studies through thorough literature reviews but also of employing meta-analytical methods to dissect the broader implications of this burgeoning academic discourse within varied educational milieus becomes paramount. Part of the novelty of what we do in this article lies in the systematicity of our approach. There are no systematic literature surveys that evaluate generative AI chatbot models within higher education, longitudinally or otherwise. Moreover, current publications on AI applications in relation to higher education still tend to be in their infancy. Efforts to establish coherence among these publications tend to be disjointed and, often, are conducted at a granular level (Ismail et al., 2023).

Familiarity with existing literature precludes inadvertent rediscovery. As a result, the following survey of the literature, available by the time of drafting this manuscript (April 2024), focuses on literature reviews and surveys that include generative AI (GenAI). Earlier chatbots (dating back to ELIZA in the 1960s) and voice-activated virtual assistants such as Siri or Alexa (in the 2010s) are, to varying extents, 'generative' (see Rudolph et al., 2023b). Whilst GenAI's most popular form in the shape of ChatGPT only burst onto the global scene in November 2022, it is preceded by foundational large language models (LLMs) and text-to-image GenAI such as DALL-E (Cao et al., 2023; Rudolph et al., 2023b). Succinctly put, GenAI can create human-like, AI-generated content, encompassing digital content such as images, music, video, and natural language (Hart, 2024; Michel-Villarreal et al., 2023).

As a consequence, there is a dearth of literature that surveys academic discussions of generative AI and higher education. Thus, for instance, Chiu et al.'s (2023) article is different from our pursuit, as it systematically reviews the opportunities and challenges of AIED by examining the

literature from 2012–2021. Similarly, Marengo et al.'s (2024) not yet peer-reviewed study has understandably little to say about GenAI and higher education as it reviewed empirical studies published between 2013 and 2022 to examine the characteristics of published research in the field of AI in higher education. Yet another example is Dogan et al. (2023), who employ a multifaceted methodological approach (integrating traditional bibliometric analysis with data mining techniques) to analyse peer-reviewed, Scopus-indexed publications that are focused on AI and written in English between 1999 to 2022. Finally, Bearman et al. (2023), while adopting a critical literature review methodology to scrutinise how AI is conceptualised within leading higher education journals, mention 'generative' AI only once in passing.

Tlili et al.'s (2023) deliberations on how AI literature reviews can be more transparent and their methodological approach is indirectly relevant to our research as it employs the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to systematically collect and evaluate 61 literature reviews on AI in education. Tlili et al. (2023) provide a detailed analysis of literature review practices in AI education research by systematically evaluating transparency through a coding scheme and identifying methodological areas needing enhancement.

While Tlili's scope is broader than ours (as it includes both K-12 education and non-generative AI), it is worthwhile noting that Stracke et al.'s (2023) study is both broader and narrower compared with our research. Stracke et al. (2023) look beyond higher education by talking about education in general terms while focusing on trustworthy and ethical AI. They introduce a unified protocol for conducting systematic reviews in AI and education (AIED), covering both the integration of AI in teaching and learning and literacy education about AI. By aligning with the PRISMA guidelines, Stracke et al.'s (2023) protocol aims to streamline research efforts, enabling consistent analysis and comparison of findings across studies. They demonstrate its utility with a review focused on trustworthy and ethical aspects of AIED, developed in tandem with the protocol to ensure mutual refinement. Stracke et al. (2023) plan to extend their innovative approach to additional key terms and extend its application over time, facilitating trend analysis and comparative research within AIED.

The above brief review shows that Ismail et al.'s (2023) observation of a dearth of systematic and macro-level research on our topic continues to be true. Our research team (based in Australia, Singapore and the UK) applied a rigorous research protocol to examine research on AI applications and higher education. In a recent protocol paper, a systematic search strategy was proposed to critically review extant research longitudinally across generative AI chatbot models within higher education (Ismail et al., 2023). Our paper applies this protocol and introduces the first version of an open-access database that systematically surveys the pertinent academic literature from November 2022 to December 2023. Our endeavour seeks to support fellow higher education researchers in gaining access to pre-examined literature on different forms of generative AI and their impact on higher education. Using a rigorous,

systematic approach, we analyse the metadata of articles published on specific types of generative AI and higher education to explore their impact on the future of higher education. In this review, the focus was on ChatGPT. By providing an open-access database (see Ismail et al., 2024), we aim to facilitate future research. In adherence to the principles of a sound and systematic review methodology, which necessitates meticulous design and execution within the bounds of established research themes (Crawford & Cifuentes-Faura, 2022), our study sets forth this research objective:

To implement a detailed research protocol designed for the systematic curation and analysis of literature on GenAI applications (such as ChatGPT), our study aims to facilitate evidence-based decision-making processes among policymakers, educators, and scholars in the higher education sector.

Consequently, our article and the resulting database employ a methodological framework intended to enable a detailed examination of the metadata and substantive findings of scholarly articles focused on GenAI applications pertinent to higher education.

## Methods

Ismail et al. (2023) provide a more detailed version of our methodical approach through an updated summary. Systematic reviews methodically compile and analyse existing knowledge within a research domain. They employ a structured approach to evaluate collective findings against predefined criteria (Higgins et al., 2011; Motyka, 2018). While research metrics serve as vital tools for assessing the quality and impact of these findings (Moed & Halevi, 2015), their inherent limitations necessitate a multifaceted evaluation approach, eschewing reliance on a single metric (Nestor et al., 2020). Our review thus selected databases based on a composite of recognised metrics, including Journal Impact Factor, h-index, g-index, Eigenfactor score, and Altmetrics, to ensure a thorough and balanced assessment of research quality (Ismail et al., 2023).

## Search strategy

Our literature survey used a systematic approach for article selection guided by PRISMA (Moher et al., 2009; Page et al., 2021). Specifically, it employed the reporting recommendations for systematic reviews suggested in the PRISMA 2020 guidelines to reflect recent developments and protocol suggestions in systematic review methodologies (see Bearman et al., 2012; Butler-Henderson et al., 2020, 2021; Page et al., 2021). Following PRISMA search guidelines, our systematic review conducted a database search of all published journal articles and preprints that relate to the topic of ChatGPT and teaching and learning in higher education.

All research outputs published between 30 November 2022 and 31 December 2023 in the following sources were considered: (1) Academic Search Ultimate, IEEE Xplore,

Informit Online, Ovid, Proquest, ScienceDirect, Scopus, and Web of Science; and (2) Google Scholar (the first ten pages for each search string were reviewed). A snowball reference analysis was also conducted based on the extracted articles. Our search strategy clearly aligned the search phrases (search terms, keywords and Boolean Operators) to the thematic dimensions relevant to the research objectives. For each search, the first core strings (higher education, artificial intelligence, and 'focal artificial intelligence') were paired with one of the other strings to complete five strings. 'Focal AI' could include reviews on diverse generative AI chatbots (e.g., ChatGPT, GPT-4, Bard/Gemini, Bing Chat, Claude, or Ernie) and generative non-chatbot AI (e.g. DALL-E, GitHub Copilot, GPT-4 plugins, Midjourney, Runway, or Synthesia), although our review focused on ChatGPT.

Table 1: Concepts, search strings and reviews guiding frames (Ismail et al., 2023, p. 58).

Concept	Search string	Review that guided this frame.
1. Higher education	"Higher education" OR university* OR college OR tertiary OR undergrad* OR graduate OR postgrad*	Butler-Henderson et al. (2022); Zawacki-Richter et al. (2019)
2. Artificial intelligence	"artificial intelligence" OR "machine intelligence" OR "intelligent support" OR "intelligent virtual reality" OR "chat bot*" OR "machine learning" OR "automated tutor" OR "personal tutor*" OR "intelligent agent*" OR "expert system" OR "neural network" OR "natural language processing"	Zawacki-Richter et al. (2019)
3. Focal artificial intelligence	ChatGPT* OR "Chat Generative Pre-trained Transformer"	<i>Use specific tool related text.</i>
4. Learning Setting	Curricul* OR learn* OR student*	Zawacki-Richter et al. (2019)
5. Education policies	Polic*	Aikens et al. (2016)
6. Assessment	Assess*	Struyven et al. (2005)
7. Teachers and lecturers	Teach* OR Lectur*	Spelt et al. (2009)
8. Pedagogical Approaches	Pegagog*	Spelt et al. (2009)

## Eligibility criteria and selection procedure

Our search was limited to English-language academic journals and pre-prints, with the review covering manuscripts published up until 31 December 2023. We included articles focusing on aspects of teaching, curriculum development, education, and student engagement in higher education, specifically those that address assessments, teaching practices, and course design related to the targeted AI tool. Exclusions were made for articles that deal with university administrative processes not pertinent to teaching or learning, as well as studies on students that do not directly relate to educational or pedagogical contexts. For instance, articles without a clear link to higher education contexts were omitted from our review (Ismail et al., 2023).

A double screening procedure was adopted in the systematic review during the verification process across the initial title and abstract screening and full-text screening to determine the final selection of sources of evidence for analysis. An appropriate reliability check (e.g., Cohen's Kappa) was conducted with at least fair agreement between all pairs required prior to progression. In the title and abstract stage, Cohen's kappa ranged between .47 and .86 across all author review pairs, except for one reviewer whose pairs were .28 and .39. These were all checked a third time for posterity ahead of progression. The quality of the evidence gathered in the systematic review was evaluated using Cochrane

Collaboration's tool for risk of bias assessment (Higgins et al., 2011; Page et al., 2021; Zeng et al., 2020) to minimise bias. The flow of information through this systematic review and aggregated findings based on the prespecified criteria was subsequently reported through a PRISMA Statement (Figure 1).

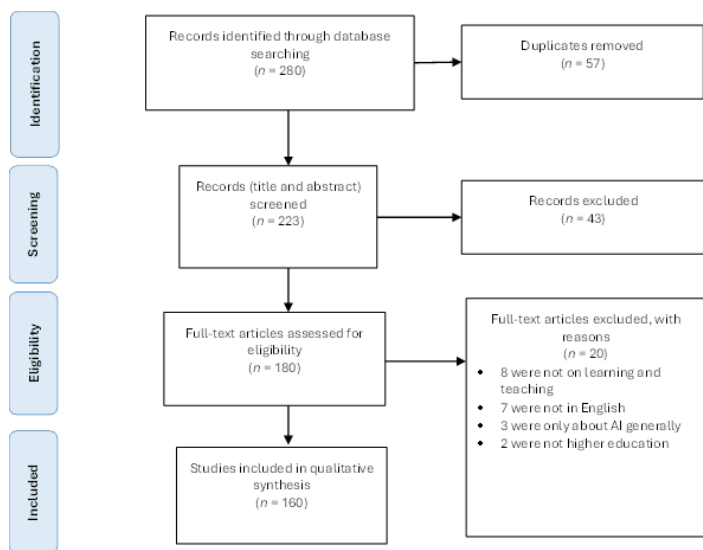


Figure 1: PRISMA statement.

### Study validity assessment

We used the PRISMA checklist and critical appraisal tools suited to the methods of the included studies to appraise and critically assess the validity of the studies (Moher et al., 2009, 2015). The PRISMA checklist is a document that guides reporting systematic reviews and meta-analyses clearly and transparently. It ensures that the systematic review is written comprehensively and transparently so that readers can assess the quality and validity of the evaluation (Page et al., 2020).

### Data coding and extraction strategy

Our data coding and extraction strategy included the production of a detailed spreadsheet that is being made available as an open-access database for scholarly reuse (Ismail et al., 2024) in conjunction with the publication of this article. In constructing the database, we incorporated certain theoretical assumptions detailed in Table 2. These are shared to present our reflexivity as researchers and to help others understand the adaptability of the data for their respective contexts. Although many data elements are clear and can be readily used in future research (like DOI, journal metadata, and country of origin), others, like the quality assessment score, study type, and participant type, necessitate further explanation.

The discipline and sub-discipline categories require some elaboration. The discipline category is grouped into four categories: health science, humanities and social science, STEM (science, technology, engineering and mathematics), and 'others' (see Butler-Henderson et al., 2020; Ismail et al., 2023). The type of study is defined as quantitative, qualitative, or mixed methods. Should there be no empirical

research, the field will be left blank. For participants, possible categories were academics, practitioners, or students – undergraduate, postgraduate and doctoral (see Butler-Henderson et al., 2020; Ismail et al., 2023).

Table 2: Description of data elements.

No.	Data Element	Field Type	Description
1.	Country of First Author	Alphabetic	Country that the Author is based in
2.	Article Type	Alphabetic	<b>The 10 categories:</b> Action research Autoethnography Case study Commentary Descriptive Opinion/Perspective Reflection Research Study Review Theoretical
3.	Type of Studies	Alphabetic	<b>The three categories:</b> Quantitative Qualitative Mixed
4.	Participant Type	Alphabetic	<b>The five categories:</b> Academics Practitioners Students Others All
5.	Discipline	Alphabetic	<b>The six categories:</b> Health Education Humanities/ Social sciences STEM Others All
6.	Title	Alphabetic	Title of the publication
7.	Author(s)	Alphabetic	Last Name, Initial of First Name of the author(s)
8.	Abstract	Alphabetic	Summary of the publication
9.	Year	Numeric	Year of Publication
10.	Journal	Numeric	Name of Journal
11.	Volume Number, Issue Number, Pages	Numeric	Details of the journal
12.	DOI/Hyperlink	URL	Digital Object Identifier / URL of the publication where available

To test the replicability of our process, the description of the above data elements was executed with different researchers. The outcomes from each repetition were recorded and compared for consistency using the metrics described in Table 2. To ensure intercoder reliability, all coders underwent standardised training using Table 2 as a shared coding manual. Their outputs were periodically cross-checked against one another to assess consistency. Reliability was statistically measured and established using Cohen's Kappa (Warrens, 2015). Conflicts in the review decision were deferred to a consensus meeting for the team to come to a resolution. This streamlined and coherent approach ensured the integrity of the database and led the team to the extraction phase of our research project.

### Results and discussion

Despite its long and rich history, AI development has made significant and noteworthy progress in the past couple of years (Haenlein & Kaplan, 2019). This includes the launch of AI-powered chatbots such as ChatGPT (Susnjak, 2022). Expectedly, the body of research examining the use of AI-



related technologies, including ChatGPT, has also expanded dramatically over the course of a year since the launch of ChatGPT-3.5 in November 2022 (Gupta et al., 2023). The geographical distribution of publications on AI can be observed through the heat map in Figure 2.

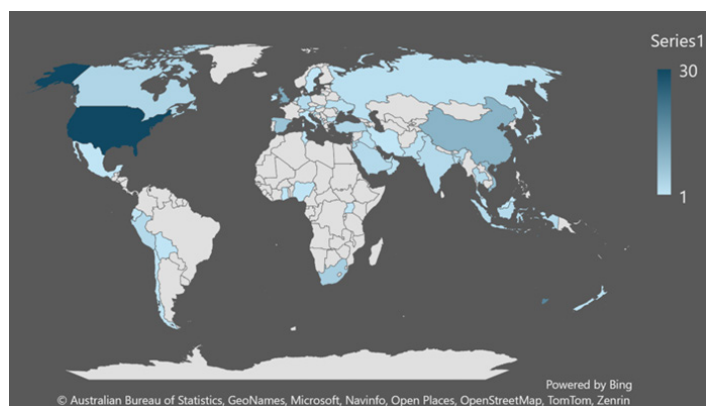


Figure 2: Heat map of geographical distribution of publications.

The first authors of the articles in our database were based in the coincidentally round number of 50 countries. The heat map indicates that the largest number of studies from a single country came from the US (28). Australia has the second-largest number of studies (18). 12 studies originated in the UK and nine in the UAE. China and Vietnam are represented with seven articles each. In terms of continents, Asia contributes 61 (38.1%), the Americas 38 (23.8%), Europe 33 (20.6%), Australasia 20 (12.5%), and Africa eight (5%) articles.

The articles include a broad range of empirical research, such as surveys, interviews, evaluations, and case studies. There were also theoretical pieces, including opinion pieces, commentaries, and reviews, as summarised in Table 3. In studies involving empirical research, qualitative studies (73; 45.6%) account for nearly half of the studies in the database, while quantitative ones (26; 16.3%) and studies using a mixed methods approach (14; 8.7%) were less popular. Nearly a third of the studies in the database were non-empirical (47; 29.4%).

The majority of the articles involved students as the primary participants (65; 40.6%) and studies having mixed groups of participants (65; 40.6%). There were fewer studies involving practitioners (7, 4.4%) and academics (18, 11.3%). Only ten studies (6.3%) were from the STEM discipline, and 14 (8.8%) were from the humanities. The 24 studies from the health discipline make up 15 per cent, but studies broadly located in education (89, 55.5%) formed more than half of the studies reviewed.

The data presented in this review provide insights into the current state of research on generative artificial intelligence in higher education. Our database offers an opportunity for research scholars to undertake future research involving AI in higher education. Given the immense potential and threats that GenAI holds for higher education, we encourage scholars to draw upon our method and database to facilitate their own research. An appropriate citation can be found in

Table 3: Summary of article characteristics.

Article Types	Numbers	%
Action research	2	1.3
Autoethnography	1	0.6
Case study	4	2.5
Commentary	8	5
Descriptive	2	1.3
Opinion/Perspective	13	8.1
Reflection	1	0.6
Research Study	89	55.6
Review	39	24.4
Theoretical	1	0.6
<b>Total</b>	<b>160</b>	<b>100%</b>

Types of study	Numbers	%
Quantitative	26	16.3
Qualitative	73	45.6
Mixed	14	8.7
Others	47	29.4
<b>Total</b>	<b>160</b>	<b>100%</b>

Participant Types	Numbers	%
Academia	18	11.3
Practitioners	7	4.4
Students	65	40.6
Others	5	3.1
All	65	40.6
<b>Total</b>	<b>160</b>	<b>100%</b>

Disciplines	Numbers	%
Health	24	15
Education	89	55.5
Humanities/ Social sciences	14	8.8
STEM	10	6.3
Others	10	6.3
All	13	8.1
<b>Total</b>	<b>160</b>	<b>100%</b>

our reference list (Ismail et al., 2024).

## Conclusion

The database attached to this manuscript provides opportunities for scholars to extract specific components of the published literature for their own studies. This database and its future versions will open the door to facilitate easy access to undertake future research based on a clear and transparent understanding of the database. We encourage scholars to download filtered versions of the database and draw on our systematic efforts in their own research (see

A note on the significance of open access (OA) publications is in order. Their growing popularity offers widespread benefits, including free and immediate access to research, enhancing its reach, impact, and efficiency, and ensuring equitable access. This stands in contrast to traditional models where taxpayer-funded research often remains inaccessible behind paywalls, a practice that limits scientific engagement (Butler-Henderson et al., 2020; Max Planck Society, 2003; Schiltz, 2018; Science Europe, 2013). Many funding bodies now mandate OA publication to ensure unrestricted access to research findings. Among OA models, Diamond OA stands out for not imposing fees on authors, thus preserving their copyright (Butler-Henderson et al., 2020; Chen & Olijhoek, 2016; cOAlition S, 2020; Fuchs & Sandoval, 2013; Olijhoek et al., 2015). The necessity of open availability of research for scientific progress is emphasised, with recent findings suggesting the value of extending open practices to data sharing (cOAlition S, 2020). We advocate this approach in our work to promote transparency but also accelerate research efforts, particularly in urgent and vital issues like AI and higher education.

Our paper details the development and research underpinning the open-access Artificial Intelligence in Higher Education Database (AIHE V1: Ismail et al., 2024). Employing a comprehensive systematic review methodology, we aimed to maximise the utility and accessibility of the data and metadata within the database. Our approach included a thorough literature review, database examination, and online resource search to encompass a wide range of publications. The process involved meticulous double-screening and double full-text review, all meticulously documented to aid fellow academics. In addition, we carefully selected and organised this database to facilitate collaboration and synergy among researchers (Butler-Henderson et al., 2020).

To the best of our knowledge, this database is the first of its kind in the higher education literature to curate the existing literature for higher education practitioners and researchers. By centralising the literature within a single database, we aim to streamline the research process, saving time for scholars while guaranteeing that a robust methodological foundation informs new studies. This convenience is anticipated to boost the production of studies exploring the diverse effects of AI on learning and teaching (see Butler-Henderson et al., 2020). Actively disseminating this resource will play a vital role in advancing the scholarship surrounding GenAI's role in education.

We will consider periodically updating and refining this methodology, incorporating future time segments, revising coding protocols, and expanding our database selection to enhance this resource's robustness and relevance over time (Ismail et al., 2023). This strategy aims better to address the effects of AI and other educational technologies, supporting the global higher education community's transition towards fresh insights in learning and teaching within the dynamically changing landscape challenged and transformed by AI applications.

## References

- Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education*, 86(2), 369-385. <https://doi.org/10.1007/s10734-022-00937-2>
- Bearman, M., Smith, C. D., Carbone, A., Slade, S., Baik, C., Hughes-Warrington, M., & Neumann, D. L. (2012). Systematic review methodology in higher education. *Higher Education Research & Development*, 31(5), 625-640. <https://doi.org/10.1080/07294360.2012.702735>
- Butler-Henderson, K., Crawford, J., Rudolph, J., Lalani, K., & Sabu, K. M. (2020). COVID-19 in Higher Education Literature Database (CHELD V1): An open access systematic literature review database with coding rules. *Journal of Applied Learning & Teaching*, 3(2), 11-16. <https://doi.org/10.37074/jalt.2020.3.2.11>
- Butler-Henderson, K., Tan, S., Lalani, K., Sabu, K. M., Kemp, T., Rudolph, J., & Crawford, J. (2021). Update of the COVID-19 higher education literature database (CHELD v2). *Journal of Applied Learning & Teaching*, 4(1), 134-137. <https://doi.org/10.37074/jalt.2021.4.1.22>
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). *A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT*. arXiv preprint arXiv:2303.04226.
- Chen, X., & Olijhoek, T. (2016). Measuring the degrees of openness of scholarly journals with the open access spectrum (OAS) evaluation tool. *Serials Review*, 42(2), 108-115. <https://doi.org/10.1080/00987913.2016.1182672>
- Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- cOAlition S. (2020). *Plan S. Making full and immediate open access a reality*. <https://www.coalition-s.org/>
- Crawford, J., & Cifuentes-Faura, J. (2022). Sustainability in higher education during the COVID-19 pandemic: A systematic review. *Sustainability*, 14(3), 1879. <https://doi.org/10.3390/su14031879>
- Dogan, M. E., Dogan, G. T., & Bozkurt, A. (2023). The use of artificial intelligence (AI) in online learning and distance education processes: A systematic review of empirical studies. *Applied Sciences*, 13(5), 3056. <http://dx.doi.org/10.3390/app13053056>
- Fuchs, C., & Sandoval, M. (2013). The diamond model of open access publishing: Why policy makers, scholars, universities, libraries, labour unions and the publishing world need to take non-commercial, non-profit open access serious. *TripleC: Communication, Capitalism & Critique*, 11(2), 428-443. <http://dx.doi.org/10.31269/vol11iss2pp428-443>

- Gupta, B., Mufti, T., Sohail, S. S., & Madsen, D. Ø. (2023). ChatGPT: A brief narrative review. *Cogent Business & Management*, 10(3). <https://doi.org/10.1080/23311975.2023.2275851>
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Hart, R. (2024, February 16). *OpenAI's Sora has rivals in the works, including from Google and Meta*. Forbes. <https://www.forbes.com/sites/roberthart/2024/02/16/openais-sora-has-rivals-in-the-works-including-from-google-and-meta/?sh=700ecc282843>
- Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., Savović, J., Schulz, K. F., Weeks, L., & Sterne, J. A. C. (2011). The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343, 1-9. <https://doi.org/10.1136/bmj.d5928>
- Ismail, F., Crawford, J., Tan, S., Rudolph, J., Tan, E., Tang, F., Seah, P., Ng, F., Visser Kaldenbach, L., Naidu, A., Stafford, V., & Kane, M. (2024). *Artificial intelligence in higher education database (AIHE V1)*. <https://doi.org/10.37074/jalt.2024.7.1.35D>
- Ismail, F., Tan, E., Rudolph, J., Crawford, J., & Tan, S. (2023). Artificial intelligence in higher education. A protocol paper for a systematic literature review. *Journal of Applied Learning and Teaching*, 6(2), 56-63. <https://doi.org/10.37074/jalt.2023.6.2.34>
- Marengo, A., Pagano, A., Pange, J., & Soomro, K. A. (2024). The educational value of artificial intelligence in higher education: A 10-year systematic literature review. *Interactive Technology and Smart Education*. <http://dx.doi.org/10.20944/preprints202311.0055.v1>
- Max Planck Society. (2003). *Berlin declaration on open access to knowledge in the sciences and humanities*. <https://openaccess.mpg.de/Berlin-Declaration>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of GenAI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://www.mdpi.com/2227-7102/13/9/856>
- Moed, H. F., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology*, 66(10), 1988-2002. <https://doi.org/10.1002/asi.23314>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264-269. <https://doi.org/10.7326/0003-4819-151-4-200908180-0013>
- Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., ... & Prisma-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4, 1-9.
- Motyka, B. (2018). Employee engagement and performance: A systematic literature review. *International Journal of Management and Economics*, 54(3), 227-244. <https://doi.org/10.2478/ijme-2018-0018>
- Nestor, M. S., Fischer, D. L., Arnold, D., Berman, B., & Del Rosso, J. Q. (2020). Rethinking the journal impact factor and publishing in the digital age. *The Journal of Clinical and Aesthetic Dermatology*, 13(1), 12. [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7028381/pdf/jcad\\_13\\_1\\_12.pdf](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7028381/pdf/jcad_13_1_12.pdf)
- Olijhoek, T., Bjørnshauge, L., & Mitchell, D. (2015). Criteria for open access and publishing. *ScienceOpen Research*. <https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AOFRQA.V1>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffman, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372. <http://dx.doi.org/10.1136/bmj.n160>
- Rasul, T., Nair, S., Kalendra, D., Robin, M., de Oliveira Santini, F., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning and Teaching*, 6(1), 41-56. <https://doi.org/10.37074/jalt.2023.6.1.29>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Rudolph, J., Tan, S., & Tan, S. (2023b). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364-389. <https://doi.org/10.37074/jalt.2023.6.1.23>
- Rudolph, J., Ismail, M. F., & Popenici, S. (2024). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. *Journal of University Teaching and Learning Practice*, 21(6). Ahead of publication.
- Schiltz, M. (2018). Science without publication paywalls: cOAlition S for the realisation of full and immediate open access. *Frontiers in Neuroscience*, 12. <https://doi.org/10.3389/fnins.2018.00656>
- Science Europe. (2013). *Science Europe principles on open access to research publications*. <http://scieur.org/opennew>
- Stracke, C. M., Chounta, I. A., Holmes, W., Tlili, A., & Bozkurt, A. (2023). A standardised PRISMA-based protocol for systematic reviews of the scientific literature on artificial intelligence and education (AI&ED). *Journal of Applied*

*Learning and Teaching*, 6(2), 64-70. <https://doi.org/10.37074/jalt.2023.6.2.38>

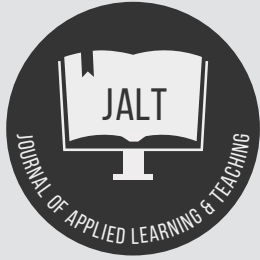
Susnjak, T. (2022). *ChatGPT: The end of online exam integrity?*. arXiv preprint arXiv:2212.09292.

Tlili, A., Huang, R., Mustafa, M. Y., Zhao, J., Bozkurt, A., Xu, L., Wang, H., Salha, S., Altinay, F., Affouneh, S., & Burgos, D.

(2023). Speaking of transparency: Are all artificial intelligence (AI) literature reviews in education transparent? *Journal of Applied Learning and Teaching*, 6(2), 44-55. <https://doi.org/10.37074/jalt.2023.6.2.15>

Warrens, M. J. (2015). Five ways to look at Cohen's kappa. *Journal of Psychology & Psychotherapy*, 5. <https://doi.org/10.4172/2161-0487.1000197>

Copyright: © 2024. Mohamed Fadhil bin Mohamed Ismail, Joseph Crawford, Shannon Tan, Jürgen Rudolph, Eunice Tan, Pauline Seah, Fiona Xiaofei Tang, Florence Ng, Laura Visser Kaldenbach, Anand Naidu, Vanessa Stafford and Michelle Kane. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



---

## Digital versus classroom discussions: Motivation and self-efficacy outcomes in speaking courses via Gather.town

---

Nimer Abusalim <sup>A</sup>	A	<i>School of Foreign Languages, The University of Jordan (Jordan)</i>
Mohammad Rayyan <sup>B</sup>	B	<i>School of Foreign Languages, The University of Jordan (Jordan)</i>
Sarah Alshanmy <sup>C</sup>	C	<i>School of Foreign Languages, The University of Jordan (Jordan)</i>
Sharif Alghazo <sup>D</sup>	D	<i>School of Foreign Languages, University of Sharjah (United Arab Emirates) &amp; School of Foreign Languages, The University of Jordan (Jordan)</i>
Ghaleb Rababah <sup>E</sup>	E	<i>School of Foreign Languages, University of Sharjah (United Arab Emirates) &amp; School of Foreign Languages, The University of Jordan (Jordan)</i>

---

### Keywords

---

Generalized self-efficacy;  
intrinsic motivation;  
metaverse;  
small group discussions;  
speaking course.

### Abstract

---

The debate on online and traditional classroom methods has gained traction in the rapidly changing world of education. This is especially true when discussing elements that are critical to the learning process, such as student motivation and self-confidence. Our research examined these components in great detail in the context of a speaking course to learn German as a foreign language. Thirty-three students made up the sample. They were divided into two groups: one that took part in traditional classroom discussions (which served as the control group) and another that experimented with discussions in the Gather.town online community. The first of our study's two goals was to determine whether the students' self-efficacy and intrinsic motivation had changed significantly after six weeks of discussion. Secondly, we were interested in determining which medium—face-to-face or online—had a stronger influence on fostering these essential educational traits. We used surveys that were distributed at the start and end of the six-week study window to gather our findings. The findings of our study were compelling. Compared to their peers in the traditional classroom environment, students who used Gather.town as their discussion platform displayed a more pronounced increase in both motivation and self-efficacy. This highlights the burgeoning potential of online learning environments like Gather.town in the modern educational landscape and suggests the benefits of incorporating such cutting-edge tools to increase student motivation and confidence.

---

### Correspondence

[m\\_rayyan@ju.edu.jo](mailto:m_rayyan@ju.edu.jo) <sup>C</sup>

---

### Article Info

Received 20 December 2023  
Received in revised form 30 January 2024  
Accepted 2 March 2024  
Available online 14 March 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.24>

## Introduction

The COVID-19 pandemic has accelerated the adoption of online and distance learning, and it is increasingly likely that the educational landscape of the future will include virtual reality or applications from the metaverse (Halasa et al., 2020). Acting as an expansive interconnected digital domain, the metaverse allows users to partake in a multitude of activities while still being linked to the real world. Conversely, virtual reality transports users into simulated realms, usually isolating them from their actual surroundings. In higher education, especially during and after the COVID-19 lockdowns, online/distance learning has become a crucial component (Abusalim et al., 2020). Applications that use the metaverse, like Gather.town, give students a virtual environment in which to interact and learn. These applications may have advantages over more conventional online and in-person learning settings. The justification behind our focus on a virtual reality mechanism simulating real-life scenarios is that Artificial Intelligence (AI) such as ChatGPT and others, have created a disruption in education habits, which may and can be beneficial if used correctly. A content analysis of news articles performed by Sullivan et al. (2023) indicated a predominant focus on academic integrity concerns and innovative assessment design in the public discourse and university responses regarding ChatGPT. To foster meaningful and intrinsically motivating learning experiences, educators are encouraged to utilize authentic assessments, which are creative learning experiences to test students' skills and knowledge in realistic situations (Rudolph et al., 2023, as cited in Wiggins, 1990). This is why we concern ourselves with one method of testing students' skills in real-life situations with the use of speaking scenarios constructed in Gather.town. Therefore, this study seeks to determine how small group discussions conducted in-person versus online using Gather.town affect students' intrinsic motivation and overall self-efficacy. This study aims to compare the psychological effects of small group discussions conducted on Gather.town versus those conducted in conventional face-to-face settings. The main goal is to identify potential differences in how well these discussion modes affect students' self-belief in their abilities. The following research questions serve as our study's compass:

1. How does the chosen medium for small group discussions, be it face-to-face or through Gather.town, impact students' sense of self-efficacy?
2. How do the modes of conducting small group discussions (in-class or online via Gather.town) affect student intrinsic motivation?

For both the initial and follow-up assessments of participants' self-efficacy levels, we will use a modified version of the Generalized Self-Efficacy questionnaire to explore these questions (Rayyan et al., 2023). Through this investigation, we hope to advance knowledge of the use of distance learning in the classroom and shed light on its psychological effects. In addition, we hope to add to the body of knowledge already available on the use of distance learning technology in education and shed light

on the psychological effects of using such technology by addressing these research questions.

In the fields of education and technology, investigating learning within metaverse applications is crucial because it represents a fundamental change in how we approach instruction and interaction. This research field is becoming more well-known because it has the potential to transform education through cutting-edge, all-encompassing, and personalized learning environments. It is crucial to look into the efficacy of metaverse applications because of the accelerated adoption of digital tools in education spurred by the COVID-19 pandemic. These tools are adaptable, useful in many areas of education, and have the potential to provide universal access to education. It is crucial to comprehend how they affect learner motivation and engagement because by doing so, educators can improve learning outcomes (Marini et al., 2022). Additionally, this research can provide beneficial pedagogical insights that will help teachers make well-informed decisions about how to incorporate metaverse applications into their teaching methods. In the end, exploring learning in metaverse applications is an essential step in preparing students for the demands of a world that is becoming more connected and digital.

## Literature review

Recent years have seen a significant increase in research on online and distance learning, particularly with the rise of COVID-19 lockdowns. As the next section will show, studies that contrast traditional classrooms with blended and fully online options have found that small group discussions have a number of advantages for improving academic performance and self-efficacy. The use of metaverse applications in education and their potential impacts on students' intrinsic motivation and self-efficacy, however, has received relatively little research.

### **Online vs. face-to-face settings: Self-efficacy and intrinsic motivation**

Education is undergoing a digital transformation that presents both opportunities and difficulties. While online learning can increase self-efficacy for some people, especially those who are tech-savvy, it might require an adjustment period for others, as we will come to see. Similarly, while flexibility and autonomy on online platforms can intrinsically motivate users, a blended strategy might enhance this motivation even further.

### ***Metaverse applications in relation to self-efficacy and intrinsic motivation***

The metaverse's emergence has sparked growing interest in its uses in education in recent years. This section looks at a number of studies that investigate how self-efficacy and intrinsic motivation are affected by metaverse applications, highlighting how important these factors are for learning.

In order to better understand how college students view the metaverse, Hwang et al. (2023) looked closely at their self-efficacy and motivation levels. The researchers discovered that students with various levels of motivation held different conceptions of the metaverse, using draw-a-picture analysis and surveys. Students with high levels of motivation tended to prefer experiential learning, which improved their growth mindsets, learning attitudes, and sense of self-efficacy. This study emphasizes how important motivation is in determining how students learn in the metaverse.

Jang and Kim (2023) focused on the effects of avatar personalization in metaverse environments, particularly in relation to fashion instruction. They looked at the impact of theoretical versus practical class modes on students' expectations and value judgments in the metaverse. Notably, it was discovered that active learning, positive expectancy, and value beliefs were enhanced by practical learning strategies. Additionally, a key factor in reducing these effects was students' creative self-efficacy. This study emphasizes how practical modes and customization can enhance learning opportunities in the metaverse.

The Technology Acceptance Model (TAM) was expanded by Al-Adwan et al. (2023) to examine the variables influencing students' intentions to adopt metaverse technology for educational purposes. Perceived usefulness, personal IT innovation, and perceived enjoyment were found in the study to be important enablers of students' behavioral intentions to adopt the metaverse. Additionally, it identified perceived cyber risk as the main deterrent. Perceived usefulness and perceived usability were found to be significantly influenced by self-efficacy, personal inventiveness, and perceived cyber risk. These results highlight the intricate interplay of variables affecting students' adoption of the metaverse.

Social cognitive theory was used by Alvarez-Risco et al. (2022) to evaluate people's intentions to engage in Facebook Metaverse activity, particularly during the COVID-19 pandemic. Their study emphasized how institutional support and technological literacy can increase one's self-efficacy for engaging in metaverse participation. Self-efficacy subsequently had a favorable impact on participants' intentions. The importance of self-efficacy and institutional support in influencing users' willingness to interact with metaverse technologies is highlighted by this study.

In his research, Choi (2022) explored the idea of immersion in metaverse applications and how it influences user engagement and recurrent use. The study examined variables like enjoyment, facilitating circumstances, and self-efficacy and found that these variables were crucial in enhancing users' immersion. A higher intention for continued use was subsequently correlated with greater immersion. This study emphasizes how crucial user involvement and immersion in metaverse experiences are.

These studies collectively shed light on the crucial connection between metaverse applications, self-efficacy, and intrinsic motivation in the context of education. They emphasize that key elements that educators and policymakers should take into account when utilizing the metaverse to improve learning environments and outcomes are motivation, customization,

institutional support, and immersive experiences.

### **Gather.town**

Gather.town is an online video conferencing platform designed for virtual conversations and business interactions in a 2D digital space. It aims to create a metaverse with human-like virtual interactions. Users receive customizable avatars for free movement, using camera, microphone, and chat functions for real-time communication. Gather is an intuitive, proximity-based video conferencing software. Users access private rooms, interact with shared documents, co-create using available objects, and connect with others. The platform offers full design features for up to twenty-five participants for free, with unlimited space creation. Although commonly used for conferences, its potential as a learning tool remains underexplored. Gather enables educators to pre-design learning spaces, communicate seamlessly between spaces and small groups, and offer tailored support in a synchronous online environment (McClure & Williams, 2021). This flexibility benefits students by fostering peer communication and a sense of identity within their learning community, while accommodating self-paced learning for developing self-regulated learning strategies (Themeli & Bougia, 2016).

To use Gather.town, educators register for a free account and choose from existing templates or customize virtual learning environments. These environments, set in scenarios like schools or hospitals, offer customizable furniture, games, and educational tools to enhance interactivity. Teachers can embed resources like videos and documents. Collaboratively created, these environments are accessible to students via a shareable link with optional password protection. Before entering, users create a customizable avatar without sign-in, choosing from various options for skin tone, hair, clothing, and accessories (Zhao & McClure, 2022).

In the 2D virtual environment of Gather.town, students interact with each other's avatars, activating a video conferencing (VC) feature in close proximity. This allows users to see, hear, and share screens. When distancing occurs, the VC feature partially disappears, simulating real-life scenarios for language practice, including everyday conversations. Students collaboratively engage with multimedia resources, such as watching videos or co-creating piano pieces. Teachers, as moderators, use a 'podium' object to broadcast instructions to all students, managing the lesson's pace. The chat function allows teachers to share links and documents either with all students or those in close proximity.

In addition to its application in virtual language lectures, Gather.town offers the unique capability of creating immersive language learning experiences. Through the platform, students have the opportunity to participate in virtual field trips to diverse global locations, providing them with an immersive encounter with the language and culture under study. This innovative approach fosters an environment where students can enhance their listening, speaking, reading, and writing skills in a dynamic and engaging manner.

Moreover, the platform serves as a conduit for interactive language immersion experiences. Students, transported virtually to different parts of the world, can submerge themselves in the language they are learning while gaining insights into the associated cultures. This immersive method not only enriches the language learning process but also provides a more dynamic context for honing language skills, contributing to a well-rounded and effective educational experience.

The use of Gather.town as a metaverse-like experience is particular to our study. This platform can be used by language teachers to create immersive language learning experiences that promote engagement and offer real-time language practice (Zhao & McClure, 2022). It improves engagement and interaction within online learning communities by incorporating gamified elements and avatars, making it a useful tool for language educators. According to a study by Latulipe and De Jaeger (2022), students preferred Gather.town over Zoom because it encouraged peer socialization, gave them more agency, and provided engaging interactions. This preference draws attention to its potential to promote group learning and raise student involvement. In addition, Gather.town's game-like environment and user-friendly features show promise in elevating engagement within higher education, thereby creating interactive virtual classrooms, according to Sriworapong et al.'s (2022) usability study.

Gather.town was successfully incorporated by Chen et al. (2022) into the educational game "Emergency Room," which was created to improve the learning process for nursing staff. The approach significantly increased learning effectiveness, according to preliminary findings, demonstrating Gather.town's potential as an effective training tool. Additionally, Gather.town's role in promoting self-paced learning in distance education was highlighted by McClure and Williams (2021), who noted that it provides unique opportunities for students to interact, customize their learning, and forge relationships in a virtual setting.

### **Self-efficacy and the mode of learning**

According to Bandura (1977a), self-efficacy refers to a person's confidence in their ability to carry out actions required to produce particular performance outcomes. It is a crucial element that fuels learning, motivation, and academic success. Artino (2012) examined students' preferences for instructional formats. According to this study, students' self-efficacy belief significantly influenced their preference for online courses. Such a tendency raises the possibility that online learning environments might give some students a feeling of control over their education, thereby boosting their self-assurance when completing academic tasks. Wang et al. (2013) emphasized the connection between technology self-efficacy and course outcomes in online learning, which lends credence to this idea. According to their research, students' self-efficacy, or confidence in their ability to use technology effectively, is a significant predictor of how well they perform academically in online environments. However, not everyone finds it easy to make the switch to online learning. Johnson (2015) compared college students

who attend on-campus and online universities. This study revealed differences in self-efficacy levels between the two groups, pointing to a potential learning curve for students transferring to virtual classrooms from traditional ones.

### **Intrinsic motivation and the mode of learning**

According to Ryan and Deci (2000), intrinsic motivation refers to the innate interest and enjoyment a person derives from a task that motivates them to complete it. This intrinsic drive can be significantly influenced by the learning style. Joo et al. (2011) investigated the factors that affect the satisfaction and perseverance of online students. According to their research, intrinsic motivation and perceived utility are the main factors that influence how satisfied online students are. Such findings suggest that, as long as students believe the content is pertinent and helpful, the autonomy and flexibility provided by online platforms can intrinsically motivate students. Carpenter and Krutka (2015) highlighted the potential of microblogging platforms like Twitter in fostering intrinsic motivation among educators in a unique examination of educators' experiences. The study demonstrated the potential of online environments in igniting and maintaining intrinsic motivation, particularly when they are interactive and community-driven, even though it did not directly compare them to face-to-face settings. However, Broadbent (2017) found that blended learners—those using both online and traditional methods—exhibited greater intrinsic motivation when comparing online and blended learners. This suggests that a hybrid learning approach, fusing the best of the virtual and physical learning worlds, might foster intrinsic motivation.

### **Small group discussions: Face-to-face vs. online**

In higher education classrooms, small group discussions have long been a popular pedagogical strategy. The development of online technologies has made it easier to use distance learning for small group discussions, particularly during and after the COVID-19 lockdowns. With the recent advancements in applications that resemble the metaverse, like Gather.town, the use of such applications in education may be in the future. Numerous studies have compared the effects of small group discussions held in person in a classroom setting versus those held remotely on student self-efficacy. In a blended learning course, Wang et al. (2019) compared the effects of in-person and online small-group discussions on students' self-efficacy. According to the study, students who participated in small group discussions online scored higher on self-efficacy tests than those who participated in in-person discussions.

Previous studies have looked into the advantages of using digital platforms in education. A study by Pellas et al. (2021) found that students felt more at ease taking part in online discussions than in conventional classroom settings. Students had favorable opinions of using digital platforms for group work, according to another study by Hernández-Sellés et al. (2019). According to this study, using digital platforms for education may increase student engagement and participation. Concerns exist, though,



regarding the negative effects that using digital platforms for education might have. According to Kim et al. (2019), online learning environments can cause students to feel lonely and disconnected. According to another study, digital platforms can make it difficult for students to collaborate and communicate effectively (Rababah, 2023).

### **Small group discussions in relation to self-efficacy and intrinsic motivation**

#### ***Small group discussions and GSE***

Small group discussions have a profound effect on students' self-efficacy, which is defined as a person's confidence in their ability to complete tasks or achieve goals (Bandura, 1977a). This effect is consistently highlighted by mainstream literature. The academic and learning paths of students are greatly influenced by their self-belief. Chang and Brickman (2018) found that participation in small group discussions led to an increase in students' confidence in their research abilities in a study looking at undergraduate research experiences.

In addition, Gokhale (1995) emphasized that college students who participated in small group discussions showed a notable improvement in their self-efficacy in relation to critical thinking. Kramarski and Mevarech (2003) found similar results, finding that students who participated in small-group metacognitive coaching and cooperative learning felt more comfortable tackling mathematical problems. In their study of the dynamics of productive cooperation in small groups, Webb et al. (1995) came to the conclusion that such a setting boosted students' self-confidence in teamwork and task completion.

According to Zimmerman and Kitsantas (2005), students who participate in regular group discussions, particularly those that are problem-solving-focused, consistently display higher self-efficacy than lone learners. In support of this, Hsiung (2013) proposed that online group discussions play a crucial role in boosting self-efficacy by providing a variety of perspectives and accessibility. Finally, Tolmie and Boyle (2000) concluded that structured peer interactions during small group discussions improved self-regulatory behaviors, which in turn increased self-efficacy.

The combined findings of the aforementioned studies highlight the transformative power of small group discussions in promoting increased student self-efficacy in a variety of educational contexts. Students learn new things and develop an innate confidence in their academic abilities through these cooperative interactions.

#### ***Small group discussions and intrinsic motivation***

It is crucial to research intrinsic motivation in educational settings. Deeper learning, greater engagement, and improved retention rates among students have all been linked to intrinsic motivation, a self-driven and inherent interest in a subject or task (Deci & Ryan, 1985). Implementing small group discussions is one pedagogical strategy that

frequently intersects with intrinsic motivation in research. Utilizing significant academic works, this literature review seeks to clarify the connection between student intrinsic motivation and small group discussions.

The Self-Determination Theory (SDT) framework, which Deci and Ryan introduced in 1985, offers fundamental insights into intrinsic motivation. The theory contends that learners exhibit intrinsic motivation when they experience autonomy, competence, and interpersonal connection during their academic endeavors. Despite not focusing solely on small group discussions, the SDT offers a theoretical framework for evaluating the efficacy of these discussions. In well-facilitated group environments where students experience ownership of their learning, share knowledge, and connect with peers. Thus, it is arguable that autonomy, competence, and relatedness are fostered.

Slavin (1996) emphasizes the idea of cooperative learning as a powerful tool for boosting intrinsic motivation by building on this. According to his research, cooperative pedagogical approaches can encourage learners' interdependence and personal accountability. Students become active participants in the ecosystem of a small group discussion, influencing and being influenced, creating a sense of shared responsibility for understanding the material. Johnson and Johnson (2009) explore this relationship in more detail by using the concepts of social interdependence. According to their research, small group activities that foster camaraderie and shared responsibility can stimulate intrinsic motivation. This idea is supported by the social interdependence theory. Discussion, debate, and analysis of ideas in a group foster a culture of collective learning, which makes tasks seem more manageable and increases individual motivation.

The effectiveness of these group dynamics, however, is not innate; rather, it depends on the makeup and circumstances of the group environment. This idea is emphasized by Cohen (1994) who contends that simply placing students in groups does not ensure success. Instead, group tasks' structured and purposeful design enhances intrinsic motivation. Students become more motivated when they understand the reason behind their discussions and can clearly see the results of their teamwork.

Järvelä and Järvenoja (2011) highlight self-regulated learning in collaborative contexts as they continue to investigate the internal dynamics of group interactions. Their findings suggest that students who are intrinsically motivated in a group setting exhibit improved self-regulation. In essence, learning motivation is honed and refined during group interactions, with peers acting as regulators and motivators for one another.

There is a longer-term perspective to consider, in addition to the immediate advantages. Hidi and Renninger (2006) first put forth the notion that situational interest, which is frequently fostered through enjoyable group activities, can develop into a significant and enduring individual interest. A topic's brief spark of interest from a single group discussion has the potential to grow into a lifelong passion or curiosity. During crucial educational phases, this ongoing interaction between group discussions and intrinsic motivation becomes

even more crucial. Cooperative learning techniques may provide a remedy for middle school students' waning intrinsic motivation, according to Anderman and Maehr (1994). Teachers may be able to rekindle the waning flame of intrinsic motivation by incorporating cooperative elements like small group discussions into the curriculum.

The importance of matching educational environments with students' intrinsic motivations is also emphasized by Lepper and Cordova (1992). Their findings point to a positive cycle in which student-driven activities, like group discussions, not only draw upon pre-existing motivations but also encourage and amplify them.

In conclusion, small group discussions are more than just a teaching strategy; they are a setting, a microcosm of the larger educational ecosystem. They act as crucibles where intrinsic motivation is nurtured and expressed when effectively structured and facilitated. Integrating pedagogical strategies that emphasize intrinsic motivation, like small group discussions, will remain essential for holistic student development as the educational landscape changes.

## Methodology

### Research design

The experimental group and the control group are two separate groups that make up the quasi-experimental research design used in this study. The control group participated in face-to-face small group discussions in a conventional classroom setting, while the experimental group participated in small group discussions via the online platform Gather.town. A pre-test and post-test using an altered Generalized Self-Efficacy questionnaire were done to assess the effect of these interventions on student self-efficacy.

### Participants

This study involved 32 undergraduate students who were all enrolled in a second-year German as a second language speaking course at the University of Jordan. 16 students were assigned to the experimental group, and 16 students were assigned to the control group, dividing the participants equally between the two groups.

### Setting

The research was done at the Faculty of Foreign Languages at the University of Jordan. The goal of the German as a second language speaking course that was provided in this environment was to improve students' language learning opportunities, with a particular emphasis on improving their oral communication abilities and competence in having real-world conversations in German. The pedagogical strategy of the course placed an emphasis on active student participation in dialogues that replicated real-life situations and were conducted entirely in German. Immersive role-playing activities were a special aspect of this course, where

students took on roles such as airline agents or passengers making German flight reservations. The curriculum also included interactive scenarios that mimicked actual places like bakeries, post offices, and movie theaters. Students engaged in dynamic and reciprocal practices within these interactive settings, ensuring exposure to both contributory and receiving sides of conversational interactions. The goal of this immersive pedagogical approach was to give students engaging language learning opportunities so they could improve their conversational competence, gain an understanding of complex cultural nuances, and hone their ability to take part in real-world German conversations.

## Instruments

### *The Generalized Self-Efficacy Scale (GSES)*

A Generalized Self-Efficacy Scale (GSES) (See Appendix 1) was used to determine the degree of student self-efficacy. Schwarzer et al. created the 10-item GSES in 1995, and it is well known for its reliability and validity. It is intended to determine how confident a person is in their ability to handle various challenging situations. This tool offers data on participants' self-confidence in their capacity to overcome obstacles and complete a range of tasks.

### *The academic intrinsic motivation questionnaire*

The subtleties of academic intrinsic and extrinsic motivation were examined by Shia (1998). Our study makes use of a modified version of Shia's (1998) academic intrinsic motivation survey (see Appendix 2). The creation of an "Academic Intrinsic Motivation" (AIM) tool to assist academic counsellors in comprehending and guiding students regarding their academic drives was a key objective of the research. The research indicates that intrinsic motivation is essential for college success. This is apparent as many students struggle with waning motivation during their time in college, leading them to seek academic advice. Shia uses Dev's (1997) definition of intrinsic motivation to highlight three key components: participation driven by curiosity, engagement for the pure joy of the activity, and the desire to contribute. Drawing on findings from Archer (1994), Miller et al. (1996), and Garcia and Pintrich (1996), the research also highlights the importance of a mastery goal.

Historical analyses draw attention to three student academic orientations: mastery, ego, and work avoidance. Using insights from Deci and Ryan (1985), Shia presents a nuanced view of intrinsic motivation, contending that it is rooted in a person's pursuit of competence and autonomy. According to Shia (1998), mastery orientation and the need for achievement are the two main components of intrinsic motivation, Shia's focus on only the "Mastery orientation" and "Need for achievement" is the result of their clear association with successful academic performance, a novel interpretation of intrinsic motivation emphasizing autonomy, and their agreement with the descriptors from the 16 Personality Questionnaire. In order to better equip academic counsellors in their advisory roles, Shia's research provides a deeper understanding of academic motivation,

highlighting the crucial roles of mastery and achievement. Therefore, in this article, we made use of only the statements pertaining to "Mastery Orientation" and the "Need for Achievement". The total number of statements was 21, given on a 7-point Likert scale (see Appendix 2).

## Procedures

Prior to the start of the Spring 2022/2023 semester, the GSES and the adapted AIM were administered to the students in both the control and experimental groups to assess the baseline self-efficacy and intrinsic motivation levels in each group. The following six weeks of small group discussions were shared by both the experimental group and the control group. Using the Gather.town platform, the experimental group held online discussions, whereas the control group held their discussions in a traditional classroom setting. In a post-test evaluation to assess any changes in self-efficacy and intrinsic motivation attributed to the interventions, each participant completed the GSES and AIM again with the appropriate adaptations.

## Data analysis

The results from the modified Generalized Self-Efficacy questionnaire were compared in order to assess any differences in self-efficacy levels between the two groups. Descriptive statistics, such as means and standard deviations, were used in the data analysis process to offer insights into the gathered data.

## Results

The findings from an investigation into the effects of small group discussions on students' intrinsic motivation and self-efficacy (conducted in-person or online using Gather.town) are presented in this section. The research questions will each be addressed separately in order to accomplish this. We begin by confirming the accuracy and dependability of the IM and GSE scales.

### Consistency validity and reliability

#### Self-efficacy scale

*Internal consistency validity.* The degree to which all of the questionnaire's items are consistent with the dimension to which they belong is referred to as the internal consistency validity of the scale's items. This indicates that each measurement only measures what it was designed to measure. The Pearson correlation coefficient was therefore calculated between each item's score and the scale's overall score. All of the correlation coefficients between the scale's individual items and total score were found to be statistically significant at a level of 0.05, ranging between 0.630 and 0.804. All of these results are statistically significant, demonstrating the Self-Efficacy Scale's high level of internal consistency. As a result, the scale's final version has 9 items.

*Reliability.* The Self-Efficacy Scale was tested for reliability using Cronbach's Alpha coefficient. The scale's overall score was a 0.86. According to the standards established by Nunnally and Bernstein (1994, pp. 264–265), who suggested 0.70 as the minimum threshold for reliability. This result suggests that the Self-Efficacy Scale has an acceptable level of reliability and can be trusted for field application.

#### Intrinsic motivation scale

*Internal consistency validity.* The Pearson correlation coefficient was calculated between the scores of each item and the scale's overall score in order to confirm the internal consistency validity of the Intrinsic Motivation Scale. At a level of 0.05, it was determined that all correlation coefficients between the scale's individual items and its overall score were statistically significant. They ranged from 0.421 to 0.867. All of these numbers are statistically significant, which suggests that the Intrinsic Motivation Scale has good internal consistency. As a result, the scale's final version has 20 items.

*Reliability.* The Intrinsic Motivation Scale's reliability was evaluated using Cronbach's Alpha coefficient. The scale's overall score was a 0.94. According to Nunnally and Bernstein's (1994, pp. 264–265) criteria, which recommended 0.70 as the minimum threshold for reliability, this value shows that the Intrinsic Motivation Scale has a high degree of reliability and can be relied upon for field application.

## Results regarding GSE

This section presents the results relating to the first research question; namely, "Are there statistically significant differences at the  $\alpha=0.05$  level between the mean scores of students in the control and experimental groups in terms of their self-efficacy, attributed to the method of teaching (conventional vs. conducting small group discussions (either in-class or online via Gather.town))?"

To answer this question, the mean scores and standard deviations of the responses from both study groups were calculated based on the pre- and post-measures of self-efficacy.

Table 1: Mean scores and standard deviations of the responses from the study groups on the pre- and post-self-efficacy measures.

Teaching Method	no	pre		post	
		Mean	Std	Mean	Std
Traditional	17	2.37	0.43	2.93	0.68
Group Discussions	16	2.40	0.38	3.51	0.24
Total	33	2.38	0.40	3.21	0.58

It can be discerned from Table 1 that there are apparent differences between the mean scores of the responses from both study groups on the self-efficacy scale, based on the group variable. The control group achieved a mean score of 2.932.93, which is lower than the mean score of the experimental group with a mean of 3.513.51. To determine whether the differences between the means are statistically

significant at the  $\alpha=0.05$  level, the Analysis of Covariance (ANCOVA) was applied. The results of the ANCOVA analysis are presented as shown in Table 2.

Table 2: Analysis of Covariance (ANCOVA) to determine the significance of the differences in the responses of the study groups on the pre- and post-measures of self-efficacy.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pre-intervention Self-Efficacy	0.551	1	0.551	2.149	0.153	0.067
Group	2.605	1	2.605	10.154	0.003*	0.253
Error	7.696	30	0.257			
Corrected Total	10.947	32				

\* The mean difference is significant at the 0.05 level.

Table 2 indicates that there are statistically significant differences at the  $\alpha=0.05$  level between the mean scores of the students on the pre- and post-self-efficacy scale according to the group variable (experimental and control). The value of F was 10.15410.154 with a significance level of 0.0030.003. To determine the effect size, the eta-squared value was calculated, which was 0.2530.253. This explains that 25.3% of the variance in student responses on the self-efficacy scale can be attributed to the group variable, while the remainder is due to other uncontrolled factors.

To determine which group had the advantage in terms of the differences in student responses on the pre- and post-measures of self-efficacy, the adjusted post-test means were extracted. Table 3 displays this information.

Table 3: Adjusted post-test means and standard errors for student responses on the self-efficacy scale.

Group	Mean	Std. Error
Control	2.94	0.12
Experimental	3.50	0.13

Table 3 reveals that the adjusted mean scores of student responses on the self-efficacy scale for the control group were 2.942.94, which is lower than the experimental group's mean score of 3.503.50. This indicates that the difference favored the experimental group, which was taught using the method of small-group discussions. These results highlight the capability of small group discussions to enhance students' self-efficacy.

### Results regarding intrinsic motivation

This section presents the results relating to the first research question; namely, "Are there statistically significant differences at the  $\alpha=0.05$  level between the mean scores of students from both the control and experimental groups in terms of intrinsic motivation attributed to the teaching method (conventional vs. small group discussions (either in-class or online via Gather.town))?"

To answer this question, the mean scores, and standard deviations of the responses from the two study groups on the pre- and post-measures of intrinsic motivation were calculated.

Table 4: Mean scores and standard deviations of the responses from both study groups on the pre- and post-measures of intrinsic motivation.

Teaching Method	no	pre		post	
		Mean	Std	Mean	Std
Traditional	17	3.14	0.85	3.86	0.85
Group Discussions	16	3.46	0.45	4.76	0.50
Total	33	3.30	0.70	4.30	0.83

It is evident from Table 4 that there are apparent differences between the mean scores of responses from the two study groups on the intrinsic motivation scale, according to the group variable. The control group (which was taught using the conventional method) had a mean score of 3.863.86, which is lower than the mean score of the experimental group (which was taught using the method of small group discussions) at 4.764.76. To ascertain if the differences between the means are statistically significant at the  $\alpha=0.05$  level, an Analysis of Covariance (ANCOVA) was employed. The results of the ANCOVA are presented in the following table.

Table 5: Analysis of Covariance (ANCOVA) to determine the significance of differences in responses from the two study groups on the pre- and post-measures of intrinsic motivation.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Pre-intervention Intrinsic Motivations	9.721	1	9.721	51.014	0.000	0.630
Group	3.332	1	3.332	17.486	0.000*	0.368
Error	5.716	30	0.191			
Corrected Total	22.212	32				

\* The mean difference is significant at the 0.05 level.

Table 5 indicates the presence of statistically significant differences at the level  $\alpha=0.05$  between the mean scores of the students on the pre-and post-intrinsic motivation scale according to the group variable (experimental and control). The F value reached 17.48617.486 at a significance level of 0.0000.000. To understand the effect size, the eta-squared value was calculated, which amounted to 0.3680.368. This explains that 36.8% of the variance in students' responses on the intrinsic motivation scale can be attributed to the group variable, while the remainder is due to other factors that are not controlled for.

To determine in whose favor the difference was concerning students' responses on the pre- and post-intrinsic motivation scale, the adjusted post-test means were extracted. The following table illustrates this.

Table 6: Adjusted post-test mean scores and standard errors of students' responses on the intrinsic motivation scale.

Group	Mean	Std. Error
Control	3.98	0.11
Experimental	4.63	0.11

From Table 6, it can be observed that the adjusted mean responses of students on the intrinsic motivation scale for the control group were 3.98. This is lower than the mean for the experimental group, which stood at 4.63. Such a difference

is indicative of the experimental group, which was exposed to the method of small group discussions, outperforming the control group. These findings underscore the efficacy of employing small group discussions in enhancing students' intrinsic motivation.

## Discussion

The landscape of education is quickly changing, clearly moving in the direction of digital platforms. This profound and undeniable shift raises important issues, particularly in relation to its impact on pedagogical practices. The focus of our study emerged from these considerations as the function and effectiveness of small group discussions conducted via different mediums, primarily face-to-face and Gather.town. Our goal was to carefully examine how these platforms affected students' self-efficacy and intrinsic motivation while grounding our conclusions in a diverse range of academic viewpoints.

Bandura's (1997a, b) theory remains a pillar of the conversation on self-efficacy. According to Bandura's theory, self-efficacy results from a person's confidence in their ability to carry out tasks. The results of our experiment were illuminating. Comparing their peers in face-to-face settings with those who participated in discussions via Gather.town, the students who did so demonstrated higher levels of intrinsic motivation. This suggests that platforms like Gather.town may be removing some restrictions, possibly the social anxieties or the numerous outside distractions that are common in conventional settings. Therefore, these virtual environments might help students feel more confident, which would increase their commitment to and participation in discussions (Stodel et al., 2006).

The framework developed by Ryan and Deci (2000) was helpful when examining intrinsic motivation. Autonomy, competence, and relatedness were identified as the three pillars supporting intrinsic motivation in their model. The manifested motivation for the cohort using Gather.town may have been a result of amplified feelings of relatedness and autonomy. Students may sense increased control and a sense of community resulting from a shared online experience within these digital boundaries. The results of Kim et al. (2019) support this viewpoint. Their study demonstrated how these virtual environments for collaboration can boost feelings of relatedness and autonomy, which in turn can spur intrinsic motivation. The groundbreaking study by Deci et al. (1999) on motivation in education is becoming ever more relevant in our technologically advanced age. Their claim that meeting students' fundamental needs can have a significant impact on their motivation raises the question of whether platforms like Gather.town are inherently more suitable for this. Bawa (2016) offers an intriguing perspective in using the compelling lens of intrinsic motivation as the key to student retention in online learning environments. Our research, which highlights Gather.town's potential advantages, can be seen as a road map for educators, instructing them on how to use digital tools to promote motivation and engagement.

Kuh (2001) contends that deep learning experiences are essential for intrinsic motivation. Online platforms offer a wealth of resources that, depending on how they are used by educators and students alike, can either promote or impede deep learning. Platforms like Gather.town can provide novelty and a novel approach to engagement, but the depth of learning experiences cultivated within these platforms may be crucial to sustaining intrinsic motivation. Although our findings demonstrate the value of platforms like Gather.town, it is critical to integrate them into a broader academic conversation. For instance, Zimmerman (2000) asserted that although social barriers on digital platforms may increase self-efficacy, they may also pose difficulties for students' ability to self-regulate. Additionally, the delicate balance between synchronous digital tools and autonomy was discussed by Giesbers et al. (2013), a factor that educators must be aware of.

In conclusion, our research reveals the complex dynamics involved in selecting the format for small group discussions. Platforms like Gather.town should not be disregarded because of their potential to affect students' intrinsic motivation and self-efficacy. Our findings highlight the need for flexibility, vision, and a dedication to utilizing the best aspects of both traditional and digital domains as the educational paradigm continues to change in response to technological advancements.

## Conclusion

The introduction of digital learning platforms has rekindled interest in comparing the relative merits of traditional and online learning strategies. Our study, which was conducted in the context of a German language course, was designed to clarify the implications of these two learning styles, particularly as they relate to group discussions and students' self-efficacy and intrinsic motivation.

Our findings demonstrated a clear benefit for the online platform Gather.town, which addressed our first research question regarding the effect of the medium on students' sense of self-efficacy. Over the course of the six-week period, the students who participated in group discussions through this medium showed a more pronounced increase in their self-belief and confidence in their language skills. In contrast, the control group showed growth but not a significant increase in self-efficacy, despite still showing growth. Moving on to our second research concern, intrinsic motivation, the digital medium once more emerged as the front-runner. Students in the Gather.town group showed increased motivation, indicating that the online setting may have provided elements that more closely matched students' intrinsic motivations. This increased motivation may have been sparked by Gather.town's freedom, adaptability, and distinctive engagement features.

The wider implications of our research must also be emphasized. Our findings support the idea that, in some circumstances, digital platforms can be more effective than conventional techniques at fostering both motivation and confidence. This is not meant to downplay the importance of in-person interactions, but rather to emphasize the potential

advantages of Gather.town-style platforms in the current educational paradigm. It calls on institutions and educators to reevaluate their pedagogical tools and methods, possibly fusing the traditional and the digital to capitalize on the advantages of both. In conclusion, our research supports the idea that digital platforms have a transformative potential for influencing educational outcomes. It is crucial for stakeholders to stay aware of these insights as education continues to follow the digital trajectory, using them to promote the best possible student development. Despite the fact that this study was limited to a particular course and platform, it opens the door for future research in a variety of fields and environments, promoting a comprehensive understanding of the digital evolution in education.

## Limitations

While our research provides insightful information about the changing nature of digital education, particularly in relation to student motivation and self-efficacy, it is important to understand its inherent limitations in order to put the results into proper perspective.

Our study's potential to be generalized is constrained by its focus on a single German language course. The findings may not be directly transferable to other courses or more general educational settings, despite the fact that they are instructive within this particular academic context. In addition, while the research's exclusive use of the Gather.town platform offers detailed insights into its effectiveness, it may not accurately reflect the effectiveness or difficulties of other digital platforms. The dynamics and features of each platform vary, so what we saw with Gather.town might not apply to another digital environment at all. Additionally, the 33-student sample size has limitations. Even though smaller samples are simpler to handle and analyze, they might miss the complex nuanced variations found in larger student populations. The six-week study period, while sufficient for our goals, may not provide a long-term view of the sustainability of the advantages associated with continued use of websites like Gather.town. Furthermore, despite the fact that the students' varied backgrounds can be extremely important, we did not go into great detail about them. Depending on their cultural, educational, or technological backgrounds, students' ability to adapt and respond to digital platforms can vary greatly, potentially biasing our findings. Finally, our reliance on surveys to measure motivation and self-efficacy poses its own set of difficulties. Even though surveys are good tools for gathering information, they are by their very nature subjective. Due to its subjectivity, self-reported data may contain biases or inaccuracies that could skew the results.

In conclusion, it is important to proceed cautiously with these conclusions even though our findings highlight the potential of websites like Gather.town in contemporary education, especially for courses like the one in our study. To develop a comprehensive understanding of the constantly changing world of digital education, more extensive research will be necessary that takes into account courses of different disciplines, larger and more diverse student populations, and multiple online platforms.

## References

- Abusalim, N., Rayyan, M., Jarrah, M., & Sharab, M. (2020). Institutional adoption of blended learning on a budget. *International Journal of Educational Management*, 34(7). <https://doi.org/10.1108/IJEM-08-2019-0326>
- Al-Adwan, A. S., Li, N., Al-Adwan, A., Abbasi, G. A., Albelbisi, N. A., & Habibi, A. (2023). Extending the technology acceptance model (TAM) to predict university students' intentions to use metaverse-based learning platforms. *Education and Information Technologies*, 28. <https://doi.org/10.1007/s10639-023-11816-3>
- Alvarez-Risco, A., Del-Aguila-Arcenales, S., Rosen, M. A., & Yáñez, J. A. (2022). Social cognitive theory to assess the intention to participate in the Facebook Metaverse by citizens in Peru during the COVID-19 pandemic. *Journal of Open Innovation: Technology, Market, and Complexity*, 8(3), 142. <http://dx.doi.org/10.3390/joitmc8030142>
- Anderman, E. M., & Maehr, M. L. (1994). Motivation and schooling in the middle grades. *Review of Educational Research*, 64(2), 287-309. <https://doi.org/10.3102/00346543064002287>
- Archer, J. (1994). Achievement goals as a measure of motivation in university students. *Contemporary Educational Psychology*, 19(4), 430-446. <https://doi.org/10.1006/CEPS.1994.1031>
- Artino A. R., Jr (2012). Academic self-efficacy: From educational theory to instructional practice. *Perspectives on Medical Education*, 1(2), 76-85. <https://doi.org/10.1007/s40037-012-0012-5>
- Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, 1(4). [https://doi.org/10.1016/0146-6402\(78\)90002-4](https://doi.org/10.1016/0146-6402(78)90002-4)
- Bandura, A. (1997b). *Self-efficacy: The exercise of control*. New York: Freeman.
- Bawa, P. (2016). Retention in online courses: Exploring issues and solutions - A literature review. *SAGE Open*, 6(1). <https://doi.org/10.1177/2158244015621777>
- Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet and Higher Education*, 33, 24-32. <https://doi.org/10.1016/j.iheduc.2017.01.004>
- Carpenter, J., & Krutka, D. G. (2015). Engagement through microblogging: Educator professional development via Twitter. *Professional Development in Education* 41(4), 707-728. <http://dx.doi.org/10.1080/19415257.2014.939294>
- Chang, Y., & Brickman, P. (2018). When group work doesn't work: Insights from students. *CBE Life Sciences Education*, 17(3). <https://doi.org/10.1187/cbe.17-09-0199>
- Chen, W. L., Ngu, P. C., & Hou, H. T. (2022). The development

- and evaluation of an online educational game integrated with Gather Town for nursing staff learning. In *Proceedings of the Asian conference on education 2021*. <http://dx.doi.org/10.22492/issn.2186-5892.2022.53>
- Choi, Y. (2022). A study on factors affecting the user experience of metaverse service. *International Journal of Information Systems in the Service Sector (IJSSSS)*, 14(1), 1-17. <http://dx.doi.org/10.4018/IJSSSS.313379>
- Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, 64(1), 1-35. <https://doi.org/10.3102/00346543064001001>
- Deci, E. L., & Ryan, R. M. (1985). Conceptualizations of intrinsic motivation and self-determination. In *Intrinsic motivation and self-determination in human behavior: Perspectives in social psychology*. Springer, Boston, MA. [https://doi.org/10.1007/978-1-4899-2271-7\\_2](https://doi.org/10.1007/978-1-4899-2271-7_2)
- Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1999). Motivation and education: The self-determination perspective. *Educational Psychologist*, 34(4), 217-230. [https://selfdeterminationtheory.org/SDT/documents/1991\\_DeciVallerandPelletierRyan\\_EP.pdf](https://selfdeterminationtheory.org/SDT/documents/1991_DeciVallerandPelletierRyan_EP.pdf)
- Dev, P. C. (1997). Intrinsic motivation and academic achievement: What does their relationship imply for the classroom teacher?. *Remedial and Special Education*, 18(1), 12-19. <https://doi.org/10.1177/074193259701800104>
- Garcia, T., & Pintrich, P. R. (1996). The effects of autonomy on motivation and performance in the college classroom. *Contemporary Educational Psychology*, 21(4), 477-486. <http://dx.doi.org/10.1006/ceps.1996.0032>
- Giesbers, B., Rienties, B., Tempelaar, D., & Gijsselaers, W. (2013). Investigating the relations between motivation, tool use, participation, and performance in an e-learning course using web-videoconferencing. *Computers in Human Behavior*, 29(1), 285-292. <https://doi.org/10.1016/j.chb.2012.09.005>
- Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education*, 7(1), 22-30. <https://doi.org/10.21061/jte.v7i1.a.2>
- Halasa, S., Abusalim, N., Rayyan, M., Constantino, R. E., Nassar, O., Amre, H., Sharab, M., & Qadri, I. (2020). Comparing student achievement in traditional learning with a combination of blended and flipped learning. *Nursing Open*, 7(4), 1129-1138. <https://doi.org/10.1002/nop2.492>
- Hernández-Sellés, N., Muñoz-Carril, P. -C., & González-Sanmamed, M. (2019). Computer-supported collaborative learning: An analysis of the relationship between interaction, emotional support and online collaborative tools. *Computers & Education*, 138, 1-12. <https://doi.org/10.1016/j.compedu.2019.04.012>
- Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, 41(2), 111-127. [https://doi.org/10.1207/s15326985ep4102\\_4](https://doi.org/10.1207/s15326985ep4102_4)
- Hsiung, C. -M. (2013). The effectiveness of cooperative learning. *Journal of Engineering Education*, 101(1), 119-137. <https://doi.org/10.1002/j.2168-9830.2012.tb00044.x>
- Hwang, G. J., Tu, Y. F., & Chu, H. C. (2023). Conceptions of the metaverse in higher education: A draw-a-picture analysis and surveys to investigate the perceptions of students with different motivation levels. *Computers & Education*, 203, 104868. <https://doi.org/10.1016/j.compedu.2023.104868>
- Jang, J., & Kim, J. (2023). Exploring the impact of avatar customization in metaverse: The role of the class mode on task engagement and expectancy-value beliefs for fashion education. *Mobile Information Systems*, 2023(1), 1-13. <http://dx.doi.org/10.1155/2023/2967579>
- Järvelä, S., & Järvenoja, H. (2011). Socially constructed self-regulated learning and motivation regulation in collaborative learning groups. *Teachers College Record*, 113(2), 350-374. <http://dx.doi.org/10.1177/016146811111300205>
- Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, 38(5), 365-379. <https://doi.org/10.3102/0013189X09339057>
- Johnson, K. (2015). Behavioral education in the 21st century. *Journal of Organizational Behavior Management*, 35(1-2), 135-150. <http://dx.doi.org/10.1080/01608061.2015.1036152>
- Joo, Y., Lim, K., & Kim, E. K. (2011). Online university students' satisfaction and persistence: Examining perceived level of presence, usefulness and ease of use as predictors in a structural model. *Computers & Education*, 57(2), 1654-1664. <https://doi.org/10.1016/j.compedu.2011.02.008>
- Kim, J., Song, H., & Luo, W. (2019). How do virtual reality platforms drive students' intrinsic motivation? *Educational Technology Research and Development*, 67(6), 1429-1448.
- Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: The effects of cooperative learning and metacognitive training. *American Educational Research Journal*, 40(1), 281-310. <https://doi.org/10.3102/00028312040001281>
- Kuh, G. D. (2001). Assessing what really matters to student learning: Inside the national survey of student engagement. *Change: The Magazine of Higher Learning*, 33(3), 10-17. <https://doi.org/10.1080/00091380109601795>
- Latulipe, C., & De Jaeger, A. (2022, February). Comparing student experiences of collaborative learning in synchronous CS1 classes in Gather.town vs. Zoom. In *Proceedings of the 53rd ACM technical symposium on computer science education* (Vol. 1, pp. 411-417). <https://doi.org/10.1145/3478431.3499383>
- Lawson, A., & Martella, A. M. (2023). Critically reflecting on the use of immersive virtual reality in educational settings: What is known and what has yet to be shown? *Journal of Applied Learning and Teaching*, 6(2), 121-133. <https://doi.org/10.1002/jalt.121>

Lepper, M. R., & Cordova, D. I. (1992). A desire to be taught: Instructional consequences of intrinsic motivation. *Motivation and Emotion, 16*(3), 187-208. <http://dx.doi.org/10.1007/BF00991651>

Marini, A., Nafisah, S., Sekaringtyas, T., Safitri, D., Lestari, I., Suntari, Y., Umasih, Sudrajat, A., & Iskandar, R. (2022). Mobile augmented reality learning media with metaverse to improve student learning outcomes in science class. *International Journal of Interactive Mobile Technologies, 16*(7). <http://dx.doi.org/10.3991/ijim.v16i07.25727>

McClure, C. D., & Williams, P. N. (2021). Gather.town: An opportunity for self-paced learning in a synchronous, distance-learning environment. *Compass: Journal of Learning and Teaching, 14*(2), 1-19. <http://dx.doi.org/10.21100/compass.v14i2.1232>

Miller, R. B., Greene, B. A., Montalvo, G. P., Ravindran, B., & Nichols, J. D. (1996). Engagement in academic work: The role of learning goals, future consequences, pleasing others, and perceived ability. *Contemporary educational psychology, 21*(4), 388-422. <http://dx.doi.org/10.1006/ceps.1996.0028>

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Pellas, N., Mystakidis, S., & Kazanidis, I. (2021). Immersive virtual reality in k-12 and higher education: A systematic review of the last decade scientific literature. *Virtual Reality, 25*, 835-861. <https://doi.org/10.1007/s10055-020-00489-9>

Rababah, L. (2023). Stepping into language mastery: Virtual reality simulations as catalysts for EFL pronunciation enhancement. *Journal of Applied Learning and Teaching, 6*(Sp. Iss. 1), 110-122. <http://dx.doi.org/10.37074/jalt.2023.6.S1.10>

Rayyan, M., Zidouni, S., Abusalim, N., & Alghazo, S. (2023). Resilience and self-efficacy in a study abroad context: A case study. *Cogent Education, 10*(1). <https://doi.org/10.1080/2331186X.2023.2199631>

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching, 6*(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67. <https://doi.org/10.1006/ceps.1999.1020>

Schwarzer, R., Jerusalem, M., Weinman, J., Wright, S., & Johnston, M. (1995). *Generalized self-efficacy scale*. [https://www.researchgate.net/publication/304930542\\_Generalized\\_Self-Efficacy\\_Scale](https://www.researchgate.net/publication/304930542_Generalized_Self-Efficacy_Scale)

Shia, R. M. (1998). *Academic intrinsic and extrinsic motivation and metacognition assessing academic intrinsic motivation: A look at student goals and personal strategy*. Wheeling Jesuit University, West Virginia.

Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology, 21*(1), 43-69. <http://dx.doi.org/10.1006/ceps.1996.0004>

Sriworapong, S., Pyae, A., Thirasawasd, A., & Keereewan, W. (2022, August). Investigating students' engagement, enjoyment, and sociability in virtual reality-based systems: A comparative usability study of Spatial.IO, Gather.town, and Zoom. In *The international conference on well-being in the information society* (pp. 140-157). Cham: Springer International Publishing. [http://dx.doi.org/10.1007/978-3-031-14832-3\\_10](http://dx.doi.org/10.1007/978-3-031-14832-3_10)

Stodel, E. J., Thompson, T. L., & MacDonald, C. J. (2006). Learners' perspectives on what is missing from online learning: Interpretations through the community of inquiry framework. *International Review of Research in Open and Distance Learning, 7*(3). <http://dx.doi.org/10.19173/irrodl.v7i3.325>

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching, 6*(1), 31-40. <http://dx.doi.org/10.37074/jalt.2023.6.1.17>

Themeli, C., & Bougia, A. (2016). Tele-proximity: Tele-community of inquiry model. Facial cues for social, cognitive, and teacher presence in distance education. *International Review of Research in Open and Distributed Learning, 17*(6), 145-163. <http://dx.doi.org/10.19173/irrodl.v17i6.2453>

Tolmie, A., & Boyle, J. (2000). Factors influencing the success of computer mediated communication (CMC) environments in university teaching: A review and case study. *Computers & Education, 34*(2), 119-140. [https://doi.org/10.1016/S0360-1315\(00\)00008-7](https://doi.org/10.1016/S0360-1315(00)00008-7)

Wang, C. -H., Shannon, D. M., & Ross, M. E. (2013). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Education, 34*(3), 302-323. <https://doi.org/10.1080/01587919.2013.835779>

Wang, N., Chen, J., Tai, M., & Zhang, J. (2019). Blended learning for Chinese university EFL learners: Learning environment and learner perceptions. *Computer Assisted Language Learning, 34*(3), 1-27. <https://doi.org/10.1080/09588221.2019.1607881>

Webb, N. M., Troper, J. D., & Fall, R. (1995). Constructive activity and learning in collaborative small groups. *Journal of Educational Psychology, 87*(3), 406. <http://dx.doi.org/10.1037/0022-0663.87.3.406>

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation, 2*(1), 2. <https://doi.org/10.7275/ffb1-mm19>

Wong, C. -H., Shannon, D. M., & Ross, M. E. (2013). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Education, 34*(3), 302-323. <http://dx.doi.org/10.1080/01587919.2013.835779>



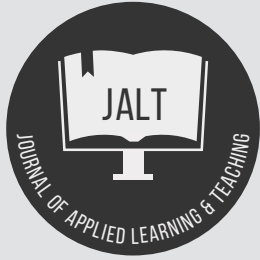
919.2013.835779

Zhao, X., & McClure, C. D. (2022). Gather.town: A gamification tool to promote engagement and establish online learning communities for language learners. *RELC Journal*, 0(0). <https://doi.org/10.1177/00336882221097216>

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology*, 25(1), 82-91. <https://doi.org/10.1006/ceps.1999.1016>

Zimmerman, B. J., & Kitsantas, A. (2005). The Hidden dimension of personal competence: Self-regulated learning and practice. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 509–526). Guilford Publications.

Copyright: © 2024. Nimer Abusalim, Mohammad Rayyan, Sarah Alshanmy, Sharif Alghazo and Ghaleb Rababah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Utilizing head simulation training in dental school education: Time and cost implications

Samuel E. Inkabi <sup>A</sup>	A	<i>College of Graduate Health Studies, A.T. Still University, Kirksville, MO, USA</i>
Isaac T. Inkabi <sup>B</sup>	B	<i>Division of Information Technology, Creighton University, Omaha, NE, USA</i>
Shirley Inkabi <sup>C</sup>	C	<i>Health Sciences Multicultural and Community Affairs Department, Creighton University, Omaha, NE, USA</i>
Adu Baffour <sup>D</sup>	D	<i>School of Science and Engineering, University of Missouri-Kansas, Kansas City, MO, USA</i>
Adwoa A. Asubonteng <sup>E</sup>	E	<i>Department of Geography and Regional Planning, University of Cape Coast, Cape Coast, Ghana</i>
Lydia E. Boamah <sup>F</sup>	F	<i>Faculty of Education, Catholic University of Ghana, Fiapre, Ghana</i>
David Benefo <sup>G</sup>	G	<i>Health Sciences Multicultural and Community Affairs Department, Creighton University, Omaha, NE, USA</i>
Philip B. Frimpong <sup>H</sup>	H	<i>Department of Oral Medicine and Pathology, University of Ghana Dental School, Accra, Ghana</i>
Obed A. Asamoah <sup>I</sup>	I	<i>School of Medicine and Life Sciences, University of Hertfordshire, Hatfield, United Kingdom</i>
Michael Oppong-Atuahene <sup>J</sup>	J	<i>Department of Medical Microbiology, University of Ghana, Accra, Ghana</i>

### Keywords

Cost;  
head simulator;  
oral health professionals;  
time.

### Abstract

Studies have suggested that the head simulator was a useful instrument for imparting hand skills for tooth removal in dental school. Although head simulator models are used by students to develop their dental hand skills, they have noteworthy limitations that restrict the breadth of knowledge and abilities that students can learn. The purpose of this current study was to determine cost and time as barriers to the effectiveness of head simulator use in dental schools. Regarding the perceived time requirements for dental courses, most participants (51.6%) disagreed that the use of head simulators extended the course duration. In terms of the availability and cost of head simulators, 40% of respondents found the availability to be above average or excellent, while 20% rated it below average or very poor. The current study suggests that the cost of head simulators did not affect the availability of the devices to oral health professionals during their school years. And the head simulator use did not extend the duration of the dental course.

### Correspondence

sa203347@atsu.edu<sup>A</sup>

### Article Info

Received 29 November 2023  
Received in revised form 10 February 2024  
Accepted 21 February 2024  
Available online 29 February 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.17>

## Introduction

Head simulators are used as experiential learning tools for dental students and oral health professionals to develop proficient dental skills (Li et al., 2021). Assessing the relationship between cost, time, and head simulator usage in dental schools is crucial to optimize resource utilization. The cost of these simulators varies widely, influenced by model type, quality, features, materials used, installation expenses, and accessory costs (Kamińska et al., 2019; Centre for Immersive Technologies et al., 2021). This cost factor is pivotal in determining their effectiveness in educational settings.

The amount of time required for students and professionals to become proficient in using head simulators is also important. It hinges on the simulator type, complexity, student training levels, and practitioners' training needs (Chernikova et al., 2020; McGleenon & Morison, 2021). Evaluating the relationship between cost, time, and head simulator utilization should be considered in their integration into health profession training. These simulators offer a realistic training environment that diminishes procedural error risks during dental practices (Li et al., 2021). Moreover, they enhance the learning experience by providing an immersive, interactive educational setting (Hamilton et al., 2021).

To examine the connection between simulator use, skill mastery, confidence levels, time investment, and associated costs, various assessment methods are viable. Using surveys and questionnaires enables an optimal evaluation of these variables among oral health professionals post-head simulator use in dental schools (Roopa & Rani, 2012). These assessment tools will gauge professionals' simulator experiences, skill levels, confidence in patient procedures, and perceptions regarding associated time and cost factors.

## Health Belief Model theoretical framework

The Health Belief Model (HBM) theoretical framework was adapted to understand the perceived barriers of cost and time in the effect that simulation training has on oral health professionals during their studies in dental institutions. The Health Belief Model (HBM) suggests that an individual's choice to engage in actions aimed at preventing or treating an illness is influenced by their perceptions of the likelihood and severity of the condition, as well as the perceived benefits and obstacles associated with its prevention or treatment (Remien et al., 2019). The exploration of the benefits and barriers of the HBM provides insight into the factors that influence educational institutions' assessments to adopt the head simulator technology against perceived barriers such as the high cost of the equipment, the lack of realism compared to real patients, and the need for additional training and time that comes with using it (Khodaveisi et al., 2021). HBM can suggest interventions that can be designed to address both the perceived benefits and barriers of head simulator training to increase participation and improve the quality of oral health care by reducing the perceived barriers of cost and time (Sanaeinasab et al., 2022).

The perceived barriers of cost and time can be explored to understand the downside of head simulator training for oral health care professionals during their studies and future engagement in this type of training. The perceived barrier of cost is a common obstacle to head simulator training. Oral healthcare professionals may perceive that the cost of the training is expensive on the part of the institution, considering the cost-benefit ratio of the training (Finocchiaro et al., 2021). A study suggested that institutions may not have access to funding for purchasing the simulator technology to integrate into the training of health professionals (San Diego et al., 2022). Utilizing the Health Belief Model's perceived barrier of cost in this study validates the concerns raised by oral healthcare professionals regarding the expenses associated with simulator training for health profession students.

The perceived time barrier is also a significant barrier to head simulator training. Solvik and Struksnes (2018) suggested that healthcare professionals perceived that they did not have enough time to participate in training because of the limited simulators available and the high demands of many students wanting hands-on practice during their clinical sessions. According to Jones et al. (2015), HBM-perceived barriers could be significant factors that prevent individuals from engaging in health-promoting behaviors. In the context of head simulator training, perceived barriers of cost and time can prevent oral health care professionals from participating in this type of training in the future, even when they recognize the potential benefits.

Medical safety is a top responsibility, and oral healthcare workers without practical experience face major consequences. Furthermore, the use of head simulators in dental schools may be hampered by time constraints. Coupled with the lower availability of head simulator technologies in dental schools, oral healthcare students may not have enough practice time with the few health simulators in their schools (Arigbede et al., 2015). In the case of appropriate availability of head simulators, it is possible that dental students already have a finite amount of practice and learning time, so adding head simulator training could detract from other crucial components of their education (Farag & Hashem, 2021).

Nonetheless, the HBM theoretical framework provides valuable insight into the potential challenges that may arise with the use of head simulator technology. Exploring the barriers of the HBM on the effectiveness of head simulators in impacting the clinical practice of oral health professionals, dental educators can promote the integration or exclusion of head simulators that will ultimately lead to improved dental education and better patient outcomes.

This current study aimed to determine oral health professionals' perceived time and cost drawbacks associated with the use of simulators during their dental studies.

## Methodology

The research project was approved by the university's Institutional Review Board, having satisfied the requirement of obtaining and submitting investigator research training certificates in human subjects' protection and financial conflict of interest (IRB number: SI20230505-001).

### Research design

A survey was conducted through Jotform (version 4.0), a secure online data collection and analysis platform, using a quantitative research approach to gather data through a purposeful sampling strategy selecting participants with the expertise to address the researcher's inquiries, specifically targeting dental hygienists, dentists, and dental assistants. The questionnaires, employing a 5-point Likert scale, included two sections, with Section 1 having 7 questions and Section 2 containing 12 questions, which investigated the barriers of cost and time associated with utilizing dental school resources. The Likert scale parameters used were "never seldom to almost always" for time, "strongly disagree to strongly agree" for cost, "very poor to excellent" for the number of available devices, and a range of time. The validation of the 5-point Likert scale was conducted by a subject matter expert and a research methodologist for reliability. The survey was constructed based on existing survey templates (Avedian, 2014). Prior to data collection, both a subject matter expert (SME) in oral health education and a research methodologist analyzed the questionnaires for content validity.

### Sample and recruitment

The population targeted for this study included dental hygienists, dentists, and dental assistants. A solicitation email was dispatched to prospective participants affiliated with the Nebraska Dental Association. Email addresses were acquired through the procurement of a mailing list from the Nebraska Department of Regulation and Licensure. The email invited recipients to voluntarily partake in the research study. The involvement of participants from the Ghana Dental Association was facilitated by sharing the research description and survey hyperlink on the association's WhatsApp platform.

The survey hyperlink directed participants to a secure data collection page on Jotform. On the Jotform platform, participants received comprehensive information regarding the study's purpose, procedures, potential risks, and benefits. Before participating in the study, written informed consent was obtained from the participant. Importantly, no personally identifiable information was collected during the study, and the data underwent anonymization during analysis. The responses from the participants were recorded for analysis and securely stored on the principal investigator's computer device, which is accessible only to the principal investigator and protected by a password.

Confidentiality was a priority and was stated in the brief description of the research study recruitment invitation that explained the proposed study and the importance of the study to health professionals' program design. Participants could opt out of the study at any time with no penalty. The survey was delivered to each participant just once to maintain data integrity and prevent participants from taking the survey more than once. Assigning unique identities, validating participant eligibility, imposing time constraints, monitoring IP addresses, and performing duplication detection during data cleaning ensured that each participant's survey was recorded only once, maintaining data accuracy and internal consistency reliability.

### Data management and analysis

To answer the research question, "What are oral health professionals' perceived time and cost drawbacks with the use of simulators during their studies?", descriptive and inferential statistics (Laerd Statistics, n.d.) were used to determine relationships between various variables. A correlation analysis was also conducted, the significance level set at  $p < 0.05$ , following Laerd Statistics (n.d.). In addition, the studied population demographics were described using measures of central tendency. The cleaned data was validated in Airtable and analyzed with the Statistical Package for the Social Sciences (SPSS, version 26).

## Results

### Demographic analysis

The ideal sample size for this study was 278, using a sample size calculator with a confidence level of 95%, a margin of error of 5%, and a population of 1000 ([www.qualtrics.com](http://www.qualtrics.com)). The survey response rate of 11.7%, with 117 responding out of a total population of 1000. Among them, 55 identified themselves as male, 61 as female, and one preferred not to disclose their gender. An analysis of the age distribution revealed that 10 individuals (8.5%) belonged to the 20 to 29 age range, indicating a substantial presence of young professionals. The age bracket of 30 to 39 years had a larger representation, with 33 participants (28.2%) falling within this category. There were 24 individuals (20.5%) aged between 40 to 49 years, signifying a significant portion of this age group. The largest segment consisted of participants aged 50 and above, accounting for 50 individuals (42.7%), highlighting the involvement of experienced professionals.

### Research findings

Out of the 117 participants, a total of 60 participants responded yes to engaging with head simulators in their dental program. The study explored oral health professionals' perceptions regarding time and cost drawbacks associated with the use of simulators during their studies. Hypotheses formulated were, the null hypothesis ( $H_0$ ) that there are no perceived time and cost drawbacks among oral health professionals who use simulators during their studies, and the alternate hypothesis ( $H_a$ ) that time and cost of

technology are perceived as major barriers in the use of head simulators for dental education.

The descriptive statistics (Table 1) analyzed responses from 60 dental program participants concerning the extension of dental course time due to head simulator use.

Table 1: Frequency distribution for profession and utilization of simulation during dental program.

Profession/Occupation	Did you use any simulation during your dental program?		Total
	No	Yes	
Dentist	42	46	88
Dental/Oral Hygienist	14	13	27
Dental Assistant	1	1	2
Total	57	60	117

In total, most respondents (51.6%, n=31) disagreed that the simulator significantly lengthened their course duration, supported by percentages across various agreement levels (see Table 2).

Table 2: Perceived extension of time requirements for dental programs due to head simulator use among dental program participants.

Rating	n	%
Strongly Agree	1	1.7
Agree	8	13.3
Neutral	20	33.3
Disagree	23	38.3
Strongly Disagree	8	13.3
Total	60	100

Participants' perceptions of the number of available head simulators concerning their cost were assessed, with the majority (40%) responding to the head simulators' availability being "Average" (Table 3).

Table 3: Perceived cost drawback for dental programs due to head simulator use among dental program participants.

Rating	n	%
Excellent	12	20.0
Above Average	12	20.0
Average	24	40.0
Below Average	8	13.3
Very Poor	4	6.7
Total	60	100

Pearson correlation analysis revealed a weak inverse relationship between extended time requirements for a course and the cost of head simulators (-0.146), suggesting insufficient evidence for a significant linear connection (Table 4).

Table 4: Correlations between dependent variables of perceived time and cost drawbacks with the use of simulators during studies.

		Extended time requirements for a course	Cost of head simulators and the number of head simulators available
		1	-0.146
Extended time requirements for a course	Pearson Correlation	--	0.266
	Sig. (2-tailed)	--	0.266
	N	60	60
Cost of head simulators and the number of head simulators available	Pearson Correlation	-0.146	1
	Sig. (2-tailed)	0.266	--
	N	60	60

Note. Time Requirements for a course and cost of head simulators: Pearson correlation: -0.146, significance (2-tailed): 0.266. Cost of Head Simulators and Number of Head Simulators Available: Pearson correlation: 1, significance (2-tailed): 0.266.

The regression analysis showed a significant relationship ( $F = 4.987, p = 0.029$ ) between the duration of simulator use during studies and the potential extension of course time (Table 5).

Table 5: ANOVA analysis of the use of the head simulator in dental schools extended the time requirements for the dental course.

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	4.195	1	4.195	4.987	0.029
Residual	48.789	58	0.841	--	--
Total	52.983	59	--	--	--

Note. A higher F-value indicates a more pronounced relationship. For this study, the F-statistic was determined to be 4.987. The significance of the relationship between the independent and dependent variables is evaluated by the associated p-value (p-value < 0.05).

The regression coefficients (Rate the amount of time spent practising with head simulators during your studies) confirmed the positive influence of simulator usage on course duration (see Table 6). This analysis supports the notion that increased time spent using head simulators during studies positively impacts the extension of course time, translating that there was no extension for the dental program duration. This provides substantial insights into the perceptions and correlations concerning simulator usage, time implications, and associated costs.

Table 6: Coefficients of the regression model on the use of the head simulator in dental schools that extended the time requirements for the dental course.

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t-value	Sig.
(Constant)	1.548	0.450	--	3.442	0.001
Rate the amount of time you spent practicing with head simulators during your studies.	0.295	0.132	0.281	2.233	0.029

Note. The constant coefficient is 1.548. This represents the expected value of the dependent variable (Simulators\_Time\_Req\_Course) when the independent variable (Time\_Training\_Simulators\_Studies) is zero.

## Discussion

In determining “What are oral health professionals’ perceived time and cost drawbacks with the use of simulators during their studies?”, most of the participants disagreed with the perception that their engagement with head simulators significantly extended the dental course time requirements. This finding suggested from the perspective of the participants that head simulators did not substantially prolong the duration of their educational programs. About 2 in 10 participants affirmed that the use of head simulators resulted in the extension of course time requirements, whereas about 7 in 10 participants strongly disagreed that the simulator had a notable effect on prolonging their course time requirements. According to Horsley and Wambach (2015), this result emphasises the argument that well-designed programs that integrate simulations enhance students’ learning experiences and skill development without significantly increasing program duration. The use of simulation improves the effectiveness of clinical training and compensates for inadequate faculty members (Horsley & Wambach, 2015).

To understand how oral health professionals who used simulators during their dental program perceived the availability of head simulators regarding their cost, participants were asked to rate this aspect. The results suggested that a substantial number of participants found the availability of head simulators to be in line with their cost expectations. Analyzing the number of participants who considered the availability of head simulators to meet the needs of oral health professional students suggested that most participants did not view the cost of head simulators as a significant drawback during their education.

The use of head simulators did not impose significant time or cost burdens during the educational pursuits of oral health professionals. This result aligns with Nabovati et al. (2022) and Rubbelke et al. (2014), who found that students can engage with simulation devices without the burden of additional expenses. In addition, the results suggest that health profession programs investing in simulators can be a cost-effective approach to health professions’ education in the long term, although the initial financial commitment in simulator technology may be substantial (Maloney & Haines, 2016). This contradicts the initial limitation suggestion that the cost associated with acquiring and upkeeping simulators, along with the investments in time and resources for training, may impose constraints on the availability of simulation-based training for practitioners (Datta et al., 2012). Therefore, the results emphasized that head simulators can reduce the time and cost required for training (Boeldt et al., 2019 ). Although initially suggested as being barriers based on the HBM theoretical framework, the current study elucidated that usage time and cost of head simulators do not pose challenges in the training of oral health professional students. This leads to comprehending the use of head simulators to assess its efficacy for oral health professional students in becoming proficient in clinical (Nabovati et al., 2022) performance.

## Conclusion and recommendation

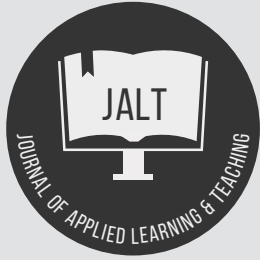
The study established, contrary to the initial hypothesis, that the use of head simulators does not lead to a substantial extension of course duration or increased financial burdens. Acknowledging the tangible benefits of integrating technology into the training of oral health professionals, this knowledge can be used by current dental educators to optimize training and better prepare dental students for the challenges they will face as practitioners.

## References

- Arigbede, A., Denloye, O., & Dosumu, O. (2015). Use of simulators in operative dental education: Experience in southern Nigeria. *African Health Sciences*, 15(1), 269–277. <https://doi.org/10.4314/ahs.v15i1.35>
- Avedian, A. (2014). *Survey design*. Harvard Law School. <https://hnmcp.law.harvard.edu/wp-content/uploads/2012/02/Arevik-Avedian-Survey-Design-PowerPoint.pdf>
- Boeldt, D., McMahon, E., McFaul, M., & Greenleaf, W. (2019). Using virtual reality exposure therapy to enhance treatment of anxiety disorders: Identifying areas of clinical adoption and potential obstacles. *Frontiers in Psychiatry*, 10, 773. <https://doi.org/10.3389/fpsy.2019.00773>
- Centre for Immersive Technologies, Mushtaq, F., & Mon-Williams, M. (2021, July 19). *Simulation-based dental education: An international consensus report*. <https://doi.org/10.31219/osf.io/c27pe>
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). *Simulation-based learning in higher education: A meta-analysis. Review of Educational Research*, 90(4), 499–541. <https://doi.org/10.3102/0034654320933544>
- Datta, R., Upadhyay, K., & Jaideep, C. (2012). Simulation and its role in medical education. *Medical Journal Armed Forces India*, 68(2), 167–172. [https://doi.org/10.1016/S0377-1237\(12\)60040-9](https://doi.org/10.1016/S0377-1237(12)60040-9)
- Farag, A., & Hashem, D. (2021). Impact of the haptic virtual reality simulator on dental students’ psychomotor skills in preclinical operative dentistry. *Clinics and Practice*, 12(1), 17–26. <https://doi.org/10.3390/clinpract12010003>
- Finocchiaro, M., Cortegoso Valdivia, P., Hernansanz, A., Marino, N., Amram, D., Casals, A., Menciassi, A., Marlicz, W., Ciuti, G. & Koulaouzidis, A. (2021). Training simulators for gastrointestinal endoscopy: Current and future perspectives. *Cancers*, 13(6), 1427. <https://doi.org/10.3390/cancers13061427>
- Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: A systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education*, 8, 1–32. <https://doi.org/10.1007/s40692-020-00169-2>

- Horsley, T. L., & Wambach, K. (2015). Effect of nursing faculty presence on students' anxiety, self-confidence, and clinical performance during a clinical simulation experience. *Clinical Simulation in Nursing*, 11(1), 4–10. <https://www.sciencedirect.com/science/article/abs/pii/S1876139914001741>
- Jones, C. L., Jensen, J. D., Scherr, C. L., Brown, N. R., Christy, K., & Weaver, J. (2015). The Health Belief Model as an explanatory framework in communication research: Exploring parallel, serial, and moderated mediation. *Health Communication*, 30(6), 566–576. <https://doi.org/10.1080/10410236.2013.873363>
- Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R. E., Avots, E., Helmi, A., Ozcinar, C., & Anbarjafari, G. (2019). Virtual reality and its applications in education: Survey. *Information*, 10(10), 318. <https://doi.org/10.3390/info10100318>
- Khodaveisi, M., Oshvandi, K., Bashirian, S., Khazaei, S., Gillespie, M., Masoumi, S. Z., & Mohammadi, F. (2021). Moral courage, moral sensitivity and safe nursing care in nurses caring of patients with COVID-19. *Nursing Open*, 8(6), 3538–3546. <https://doi.org/10.1002/nop2.903>
- Laerd Statistics. (n.d.). *Descriptive and inferential statistics*. <https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php>
- Li, Y., Ye, H., Ye, F., Liu, Y., Lv, L., Zhang, P., Zhang, X., & Zhou, Y. (2021). The current situation and future prospects of simulators in dental education. *Journal of Medical Internet Research*, 23(4), e23635. <https://doi.org/10.2196/23635>
- Maloney, S., & Haines, T. (2016). Issues of cost-benefit and cost-effectiveness for simulation in health professions education. *Advances in Simulation (London, England)*, 1, 13. <https://doi.org/10.1186/s41077-016-0020-3>
- McGleenon, E., & Morison, S. (2021). Preparing dental students for independent practice: A scoping review of methods and trends in undergraduate clinical skills teaching in the UK and Ireland. *British Dental Journal*, 230, 39–45. <https://doi.org/10.1038/s41415-020-2505-7>
- Nabovati, E., Jeddi, F. R., Ghaffari, F., & Mirhoseini, F. (2022). The effects of simulation training on learning of health information systems: A scoping review. *Journal of Education and Health Promotion*, 11, 4. <https://pubmed.ncbi.nlm.nih.gov/35281403/>
- Remien, R. H., Stirratt, M. J., Nguyen, N., Robbins, R. N., Pala, A. N., & Mellins, C. A. (2019). Mental health and HIV/AIDS: The need for an integrated response. *AIDS (London, England)*, 33(9), 1411–1420. <https://doi.org/10.1097/QAD.0000000000002227>
- Roopa, S., & Rani, M. S. (2012). Questionnaire designing for a survey. *Journal of Indian Orthodontic Society*, 46(4), 273–277. [https://www.researchgate.net/publication/235801675\\_Questionnaire\\_Designing\\_for\\_a\\_Survey](https://www.researchgate.net/publication/235801675_Questionnaire_Designing_for_a_Survey)
- Rubbelke, C. S., Keenan, S. C., & Haycraft, L. L. (2014). An interactive simulated electronic health record using Google Drive. *Computers, Informatics, Nursing: CIN*, 32(1), 1–6. <https://doi.org/10.1097/CIN.0000000000000043>
- San Diego, J. P., Newton, T. J., Sagoo, A. K., Aston, T. A., Banerjee, A., Quinn, B. F. A., & Cox, M. J. (2022). Learning clinical skills using haptic vs. phantom head dental chair simulators in removal of artificial caries: Cluster-randomized trials with two cohorts' cavity preparation. *Dentistry Journal*, 10(11), 198. <https://doi.org/10.3390/dj10110198>
- Sanaeinasab, H., Saffari, M., Taghavi, H., Karimi Zarchi, A., Rahmati, F., Al Zaben, F., & Koenig, H. G. (2022). An educational intervention using the health belief model for improvement of oral health behavior in grade-schoolers: A randomized controlled trial. *BMC Oral Health*, 22(1), 94. <https://doi.org/10.1186/s12903-022-02132-2>
- Solvik, E., & Struksnes, S. (2018). Training nursing skills: A quantitative study of nursing students' experiences before and after clinical practice. *Nursing Research and Practice*, 8984028. <https://doi.org/10.1155/2018/8984028>

Copyright: © 2024. Samuel E. Inkabi, Isaac T. Inkabi, Shirley Inkabi, Adu Baffour, Adwoa A. Asubonteng, Lydia E. Boamah, David Benefo, Philip B. Frimpong, Obed A. Asamoah, and Michael Oppong-Atuahene. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Self-perception of university teachers on their digital teaching competence: The case of Peru

---

Mercedes Rojas-Osorio <sup>A</sup>	A	<i>Escuela de Posgrado, Universidad San Ignacio de Loyola</i>
Shyla Del-Aguila-Arcenales <sup>B</sup>	B	<i>Escuela de Posgrado, Universidad San Ignacio de Loyola</i>
Aldo Alvarez-Risco <sup>C</sup>	C	<i>Universidad Tecnológica del Perú</i>

---

### Keywords

Digital competence;  
education;  
Peru;  
self-perception;  
university teachers.

---

### Abstract

Research on university professors' digital competence is essential because the university professor is the most valuable human resource educational institutions have and is responsible for carrying out various actions to achieve institutional purposes. The present study aimed to analyze university professors' self-perception of digital competence. It is a quantitative, non-experimental, descriptive, and cross-sectional study. The sample was 122 professors from one private university in Peru. Also, a significant relationship was found between gender and knowledge of methodological strategies for networking. The teacher's age was the most relevant factor in digital competence, affecting teachers' ability to know and use communication and information tools. A significant percentage of teachers did not participate in training activities and did not evaluate their teaching practices with ICT; likewise, there was low participation in projects and groups to innovate and research teaching topics with ICT. Based on the deficiencies found, a continuous training program is expected to be designed to raise teachers' digital competence levels.

### Correspondence

[sdelaguila@usil.edu.pe](mailto:sdelaguila@usil.edu.pe)<sup>B</sup>

---

### Article Info

Received 13 October 2023  
Received in revised form 1 December 2023  
Accepted 11 January 2024  
Available online 16 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.8>



## Introduction

The current context is characterized by the use of the Internet, the development of digital media, and the revolution of various concepts such as interactivity, connection, linking, and empowerment of information technology. The growth in the use of the Internet is mainly in the behaviour of the population through online shopping (Rybczewska & Sparks, 2022; Soegoto & Eliana, 2018), daily communications (Pekkala & van Zoonen, 2022), and education (González-Gutierrez et al., 2022; Hamadi et al., 2022; Rojas-Osorio & Alvarez-Risco, 2019). The change in Internet usage patterns has forced universities to update their teaching methods (Potter et al., 2022) and to develop digital competencies among professors (Amhag et al., 2019). Since 2020, due to the COVID-19 pandemic in Peru, there have been several educational reforms to ensure the continuity of education through digital media (Gobierno del Perú, 2020; MINEDU, 2020b), which has forced better levels of digital competence in the teaching staff at the university level. Therefore, it is essential to have quality university teachers because they are the ones who train future professionals in society. Several researchers emphasize the importance of including digital competence in teaching to contribute to student learning through ICT (De la Calle et al., 2021; Fernández-Batanero et al., 2020; Garzón Artacho et al., 2020), which has originated the need for educational institutions to organize training plans in technology to help their teachers acquire or strengthen their digital competence (Cabero-Almenara et al., 2020; Howard et al., 2021). However, some still cannot use ICT during the teaching and learning process.

Previously, international organizations have pointed out that university education must expand the training of teachers so that they can have the capacity to provide their students with the knowledge and skills they need in the 21st century to use distance learning and information and communication technologies (UNESCO, 2009). UNESCO recommends encouraging the Member States to (a) develop and implement ICT-enriched pedagogies, (b) make ICT skills a fundamental component of the teaching profession for all teachers at all levels, (c) support teachers in integrating ICT into all areas of their professional activity actively (teaching, research), (d) implement ICT-enriched pedagogies in teacher education programs and institutions, (e) ensure that teachers have access to ICT tools and materials and professional networks; (f) disseminate the UNESCO ICT Competency Framework for Teachers (ICT-CFT), as well as other relevant UNESCO publications, to appropriate institutions and agencies" (UNESCO, 2015).

In Peru, higher education institutions are divided into universities and higher education schools. Universities can be public or private, while higher education schools are generally focused on technical and technological training. In turn, universities can offer undergraduate and graduate programs in various disciplines. According to the National Superintendency of Higher University Education (SUNEDU for its acronym in Spanish), Peru has 50 public and 92 private universities (SUNEDU, 2023). Of 1,423,731 university students, 25.5% study in a public university and 74.5% in a private university (MINEDU, 2023).

One of the biggest challenges faced by Peruvian universities is the transformation of face-to-face teaching into a virtual modality, which saw its starting point in 2003 with the general education law describing distance education. An essential aspect of the law is that "this modality aims to complement, reinforce or replace face-to-face education, attending to the needs and requirements of the people". This replacement was seen during the COVID-19 pandemic. This change highlighted the need to train university professors in digital competencies, mainly in topics related to the knowledge and use of ICT tools for implementing resources, activities, and videoconferencing rooms. This training allows the development of classes in their various modalities, as well as design guidelines and strategies for teaching, tutoring, and student support using technology, which is linked to the three central guidelines provided by the Ministry of Education in the numerals 5.1 Academic planning, 5.2 Development of service delivery and 5.3 Recommendations related to the teaching exercise regarding the non-face-to-face adaptation (MINEDU, 2020a). As a result of those mentioned above, the Ministry of Education (2021) approved the technical norm N°109-2021-MINEDU to implement the "Permanent Training and Education Program during the year 2021", whose purpose is to strengthen the professional competencies of teachers, the training offer prioritizes four lines: learning, digital competence, research and innovation and socioemotional well-being (MINEDU, 2021).

The COVID-19 pandemic accelerated the digital transformation. The transformation in Peru has been heterogeneous, as Internet connectivity varies considerably, being more significant in the capital, Lima, and in the big cities. However, due to Peru's geography, there are areas in the highlands and jungle where connectivity was limited, which meant that during the pandemic, classes were halted. In its place, the Peruvian state channel broadcasts a primary education program to try to provide a solution. On the other hand, adaptation was successfully taking place in those large cities, using systems such as Blackboard, Canvas and Meet for synchronous classes and asynchronous accompaniment. However, compared to other countries, the process was more difficult due to connectivity, as Peru ranks 71st in Average Internet Speed, 60th in Median Internet Speed and 75th in Average Mobile Speed (Wisevoter, 2023).

In the face of these forced changes, the research question arises: What is the perception of university faculty about their digital teaching competence?

## Theoretical framework

Teachers' use of ICT involves mastering a series of competencies; it is not just using some tools. It is necessary to have a teaching model or methodology that justifies how to incorporate and use ICT in the educational process (Suárez-Rodríguez et al., 2018). Nowadays, there are models of teaching and learning where ICT is included, which has originated training needs in specific competencies to develop within the educational field known as digital teaching competence.

## Constructivism in education

The theory of constructivism is an educational approach that maintains that learning is an active process in which students construct their knowledge through interaction with information and their environment (DeVries, 2000). This theory suggests that students are not passive recipients of information but are active participants in constructing their understanding of the world. When applied to remote classes, constructivism remains relevant and can have significant implications for how learning experiences are designed and carried out in virtual environments. Here are some critical connections between constructivism theory and distance learning.

Constructivism emphasizes the importance of students actively participating in their learning. In distance classes, students often have more control over their study time and pace. Online course design can encourage self-direction by providing varied resources and activities that allow students to explore and construct their understanding independently. On the other hand, constructivism highlights the importance of social interaction in learning (Feyzi Behnagh & Yasrebi, 2020). Despite the physical distance in online classes, tools and platforms can be incorporated to facilitate student collaboration. Discussion forums, videoconferences, and online group projects can promote the social construction of knowledge.

Likewise, constructivism emphasizes the importance of connecting new knowledge with students' previous experiences. In distance learning environments, case studies, practical situations and relevant examples can contextualize information and make it meaningful to students, thus facilitating the active construction of knowledge. Also, constructivism advocates continuous feedback as an integral part of the learning process (Chuang, 2021). In remote classes, educators can leverage online tools to provide regular formative feedback to students, helping them reflect on their learning and adjust their understanding.

Finally, constructivism recognizes the diversity of learning styles and the importance of using various resources. In distance classes, multimedia resources, such as videos, simulations, and interactive activities, can be incorporated to address different learning styles and enrich the learning experience.

## Competence

The term *competence* appears in Latin in the form of *competitio* with the meaning of being able and in the form of *competens*, the understanding of which is capacity and allowance. Throughout history, the first use of the term *competencies* is found in the work of Plato, whose origin of the word is *ikano*, a derivative of *iknoumai*, which means to arrive (Miró Vera, 2019). Likewise, in ancient Greek, its equivalent for *competence* was *ikanótis*, which translates as the quality of being *ikanos* (capable), i.e., having the skill, ability, and capacity to achieve something (Mulder et al., 2008). This term is also like *Epangelmatikes ikanotita*, which means professional/vocational ability or competence.

It should also be mentioned that in the 16th century, the concept was already used in French, English, and Dutch; at the same time, the use of the words *competence* and *competency* in Western Europe dates back to the same period, so it is clear that it has been an aspiration throughout history to be a competent and capable professional (Mulder et al., 2008).

As Naranjo et al. (2017) point out, competencies began their academic applicability in the 1960s, thanks to Chomsky, who, in aspects of the theory of syntax (1971), defines them as linguistic competence, which is put into action through communicative performance. Likewise, several authors agree that the concept of competence began in 1970 with the contributions of Chomsky's linguistics and Skinner's behavioral psychology (Trujillo-Segoviano, 2014). For Skinner (1981), competence is a behavioral model based on behavior that can be observable and verifiable, which today is called the competence approach and which is applied in the "management of human talent in organizations" as the key to the competitiveness of their workers. There are differences in the concepts of both authors; Chomsky's approaches are used in the educational process, while in the labor context, Skinner's concepts are applied to achieve productivity in organizations, relying on performance efficiency and fulfilling assigned tasks.

In the seventies, we began to talk about business and educational competencies with McClelland (1973), who points out that competencies are people's abilities to do something well. In contrast to this author, according to Palan (2003), competencies represent academic evidence. Likewise, in the 1980s, the concept of competencies in professional training (Galdeano Bienzobas & Valiente Barderas, 2010) was introduced, strengthening the construction of academic programs according to the demands of the professional profiles of the new context. Subsequently, Gagné and Briggs (1974) included the term *performance indicators* to differentiate whether a worker was competent or not, while Gardner (1982) considered competence as brain function capacities or skills of a type of intelligence (Habermas, 1984, 1987).

## Teaching competence

Bearing in mind that competencies are aptitudes or capacities to carry out a specific desired performance, Bunk (1994) points out that professional competencies can be assumed from two fronts: on the one hand, there are the formal competencies that are acquired during training, and on the other hand, there are the fundamental competencies that are related to the ability to solve specific problems. In other words, the formal competencies and competencies acquired through experience are evaluated. Therefore, teaching competencies are the knowledge, skills, and attitudes to exercise a profession, solve problems autonomously (Bunk, 1994), and collaborate in their professional environment.

Aylett and Gregory (1997) establish two types of criteria: first, the criteria of competence of the teaching function referred to organizing and presenting academic information, establishing social or interpersonal relationships, being

willing to provide support or guidance to the student, and evaluating the evidence of learning and second, the criteria of excellence refer to reflecting on their teaching work, innovating their teaching-learning strategies, designing and participating in the curriculum of courses or programs, researching aspects related to teaching, organizing and executing courses, as well as leading groups and work teams with teachers. UNESCO (1998) establishes a series of competencies according to the members of the university environment: academic personnel, administrative personnel, managers, and institutional leaders. About the academic staff, the faculty is requested to use diverse methods to teach and include new technologies, that is to say, to know the training processes supported using technology.

Regarding the research actions of university teachers, they are related to the competencies of knowing how to write research projects, raise funds, participate in professional research networks, advise students and manage research projects (Böttcher-Oschmann et al., 2021; Castillo-Martínez & Ramírez-Montoya, 2021). Perrenoud (2004), and Zabalza and Beraza (2003) point out that teaching competence refers to aspects related to knowledge and skills whose actions of mobilizing these cognitive resources allow reaching required results, achieving objectives, carrying out an activity, and solving a problem in each context. Therefore, referring to teaching competence implies considering the values, beliefs, knowledge, skills and attitudes the teacher must have for the teaching-learning process, the educational institution, the needs of the subjects who learn, and professional ethics. An important aspect to highlight is that Perrenoud (2004) details ten domains that should be considered as priorities in teacher training, such as organizing and encouraging learning situations, involving students in their learning and their work, working in teams, participating in school management, and informing and involving parents.

Muñoz (2006) points out the need for a competency-based plan for the profession to professionalize teachers at the different levels where they teach. Likewise, the author provides several alternatives for a new and different teaching profession, such as collaborating with other professionals, increasing communication among professionals, receiving theory so as not to fall into reproductive practices, increasing awareness that teaching and learning are complex, that teaching is imbued with much diversity and training in introducing new technologies in teaching.

### **Dimensions of teaching competence**

University teachers require training so that they can learn and teach competencies. A teacher can carry out his professional work considering aspects related to the training process, as well as having the vocation of service as a teaching professional to teach well and with the desire to do things well for the benefit of his students. Hence, Zabalza (2009) proposes the following dimensions of competence: a) Ability to plan the teaching/learning process; b) ability to select and present the disciplinary contents; c) ability to inform and explain understandably; d) technological literacy; e) ability to manage methodological strategies and learning activities; f) ability to foster a constructive relationship

and a good atmosphere in the classroom; g) tutoring and accompanying of students; h) ability to reflect on their teaching practice and investigate to improve this process; i) ability to get institutionally involved.

According to OECD (2009), the teacher must manage learning processes, teach in multicultural classrooms, integrate students with special needs, work well in a team, plan evaluations, manage ICT, collaborate on projects, and communicate effectively with parents to build a community for learning. In the same way, the European Commission (2010) mentions that effective educators must have strong teamwork skills, be proficient in leveraging technology for teaching, and be active contributors to the educational community. Also, the European Commission (2012) proposes a list of competencies grouped into some dimensions. Competencies in this area include knowledge of the subject, pedagogy, curriculum, educational policies, inclusion, use of ICT, teaching and evaluation methods, planning, managing, teaching, monitoring, achieving learning objectives, research, collaboration, and adaptability. Dispositions include a predisposition to change, continuous learning, commitment to promoting learning, democratic practices and attitudes, and critical evaluation of oneself as a teacher.

Developing competencies in the educational field covers a wide range of skills and knowledge essential for teaching effectiveness. These competencies encompass a deep understanding of content, pedagogy, curriculum and educational policies. A competent teacher not only masters his subject(s) but also understands effective pedagogical strategies to convey that knowledge in a clear and accessible way to students. Inclusion is highlighted as a crucial competency, referring to educators' ability to adapt and attend to diversity in the classroom. It involves creating inclusive environments that allow all students to participate and learn meaningfully, regardless of their particular abilities or characteristics. Furthermore, the effective use of Information and Communication Technologies (ICT) is considered an essential competency in today's digital age, as it can significantly improve teaching and learning.

Other key competencies include teaching and assessment methods, efficient planning and management of time and resources, and the ability to monitor and adapt instruction according to the changing needs of students. Achieving learning objectives is essential, and this involves setting clear goals, constantly evaluating progress and adjusting strategies accordingly. Research is presented as a valuable competence since educators must be up to date on educational trends and commit to continuous improvement. Collaboration with other educators and the community is essential to enrich teaching practice and foster a collaborative learning environment.

Regarding the dispositions, fundamental attitudes that contribute to professional development and success in the classroom stand out. Predisposition to change reflects the ability to adapt to new educational methodologies and approaches, which is essential in a constantly evolving environment. Continually pursuing learning reinforces the importance of staying up to date and committed to personal and professional development. Dedication to promoting

learning highlights the educator's responsibility to inspire students and foster an environment that stimulates curiosity and inquiry. Promoting democratic practices and attitudes suggests a commitment to equity, participation and social inclusion in the educational environment. Critical evaluation of oneself as a teacher comes to a close of the list of dispositions, highlighting the importance of continuous reflection on practice and the willingness to adjust approaches and strategies based on the observed results.

The Finnish Institute for Educational Research (2010) reveals eight competencies: Collaboration and cooperative learning; effective teaching methods; integration of theory and practice; research-based learning; knowledge management, strategy implementation; leadership promotion; lifelong learning preparation; mobility and cultural understanding' and quality evaluation and improvement. These competencies cover various aspects crucial for the effective development of educational systems. Firstly, it highlights the importance of collaboration and cooperative learning. This approach encourages student interaction, promoting an environment where the exchange of ideas and joint problem-solving are fundamental to the learning process. According to the report, the effectiveness of teaching methods constitutes another essential pillar. It recognizes the need for effective pedagogical strategies that transmit knowledge and stimulate students' interest and active participation. Integrating theory and practice, supported by research-based learning, is an essential component of a comprehensive educational approach to develop practical skills.

Knowledge management and strategy implementation feature prominently in the report. The ability to organize and apply knowledge effectively is considered crucial to educational success. Likewise, the importance of promoting leadership in the educational field is emphasized, recognizing that educational leaders play a fundamental role in the direction and development of educational institutions.

Continuous learning preparation is a critical competency in a constantly evolving world. The ability to adapt to new knowledge and contexts is presented as essential for lifelong education. Mobility and cultural understanding are also highlighted, underlining the importance of openness to diverse perspectives and experiences in an increasingly globalized world. Quality evaluation and improvement close the list of competencies outlined in the report. The ability to evaluate the effectiveness of educational methods and make continuous improvements is essential to guarantee a dynamic educational system adjusted to the changing needs of society.

### **Teaching digital competence**

Several authors agree that the opportunities offered by ICT-mediated environments for learning enhance the role of the teacher, who facilitates and guides the search for and processing of information, making it possible for the information provided to become knowledge (Agyei, 2021; Alt, 2018; Atmacasoy & Aksu, 2018; Qaddumi et al., 2021). Similarly, some authors (Antonietti et al., 2022; Falloon, 2020;

Pöntinen & Rätty-Záborszky, 2020; Pozo-Sánchez et al., 2020; Skantz-Åberg et al., 2022) agree in defining digital teaching competence as the ability to perform teaching functions about the following aspects: a) Technical b) Academic or pedagogical c) Planning and organization d) Guidance e) Social.

Yot Dominguez and Marcelo (2005) define digital teaching competence as skills to manage and use technological resources necessary for designing and developing e-learning, referring to knowing how to use synchronous and asynchronous communication tools, among others. It also includes knowledge regarding the platform on which the training activity is developed to adapt it to the type of students and the course.

According to the Parlamento Europeo (2007), digital competence must be integrated at all levels of education, which is why it must be ensured that "teachers are trained in the use of ICT for teaching", better known as digital teaching competence. Bustos López and Gómez Zermeño (2018) identified the dimensions of teachers' digital competence in a high school in Mexico through surveys of 12 teachers and interviews with four teachers. The data was analyzed based on three categories: digital competencies, use of educational technology, and social and professional development. The outcomes showed some chances for teachers as the design resources with advanced technology to create and publish and share material in virtual learning spaces.

Teachers need training that is more consistent with their teaching needs. Using a qualitative approach (Rossi Cordero & Barajas Frutos, 2018) a case study was developed in two schools, identifying enablers and barriers in teaching the acquisition of digital competence. The outcomes show that acquiring digital competence is the most important difficulty for teachers linked to the management of information, the didactic use of ICTs, the evaluation of digitally mediated activities, and collaboration in virtual environments.

In Latin America, the various advances in the digitalization of education in times of COVID-19 have been reported, as in the case of Argentina (Perrotta, 2021), which presents the efforts in public universities for teaching and learning, research and internationalization activities. There are also reports of the cultural adaptation that has had to be generated in Brazil, considering social distancing at the beginning of the pandemic (Ivenicki, 2021; Prata-Linhares et al., 2020). In Chile, it was reported that this drastic digitalization process due to the pandemic ultimately benefited their learning due to the challenge of rapid adaptation (Sepulveda-Escobar & Morrison, 2020).

Then, it is striking what was reported in Colombia regarding distance teaching of Anatomy, where students had to learn in a very different format from the conventional one but were able to improve their learning through authentic anatomical images. It is a great adaptation process to continue learning (Martinez et al., 2022). Something similar happened in Ecuador, where engineering courses had high student attendance. It should be noted that these good results were not random but corresponded to the result of a specific program designed during the pandemic aimed at

student engagement (Ricaurte et al., 2022).

Despite the excellent adaptation of students in Latin America, it has also been reported that many had difficulties, such as chemistry students in Mexico, who reported that 78.2% had technological issues and could not handle technology tools (Chávez-Miyauchi et al., 2021). In Brazil, 66.7% of students dropped out of at least one course in their curriculum (De la Fuente et al., 2021). Likewise, in the Dominican Republic, 67.1% of students expressed a lack of adequate technological equipment and Internet access. They felt uncomfortable studying from home (Santos et al., 2021). In Mexico, students reported unstable Internet access, which caused them to miss some sessions (Zúñiga Rodríguez & Cáceres Mesa, 2021).

## Methodology

### Sample and collection of data

The population consisted of 179 undergraduate and postgraduate teachers in regular online courses from a private university in Peru. The formula for calculating the sample size with 95% reliability and a 5% margin of error was applied, obtained by a sample of 122 teachers. The information was obtained in July 2021 by an online Google Forms questionnaire. At the beginning of the questionnaire, participants were informed that their participation was anonymous and voluntary and that the data would be used exclusively for research purposes.

### Measures

The questionnaire that was developed evaluated the self-perception of digital competence, including a few specific dimensions. The first section of the questionnaire includes the presentation of the study and sociodemographic data, time dedicated to classes (partial or complete time), professional category as a professor (contracted or ordinary), and level of education (undergraduate and postgraduate). Subsequently, 53 items were presented that make up the body of the questionnaire, which were distributed as follows:

- a. General knowledge of the computer and ICT possibilities (Items 1-6)
- b. Knowledge and use of methodological strategies for networking (Item 7)
- c. Opportunities and limitations of ICT in the teaching-learning process (Items 8-9)
- d. Importance that the teacher gives to the factors when choosing an ICT resource for the classroom (Items 10-18)
- e. Knowledge and use of the virtual campus (Item 19)
- f. Procedures for publishing educational material/scientific production on the web (Items 20, 21, 25,

27, and 35)

- g. Teacher training and evaluation of teaching practice with ICT (Items 23, 28, and 29)
- h. Use of ICT for teaching tasks (Items 22 and 24)
- i. Use of various protection measures for your equipment (Items 30-32)
- j. Autonomy for technical problems regarding your computer equipment (Items 33 and 34)
- k. Research and educational innovation (Items 36 and 37)
- l. Strategies to promote student participation (Items 38 and 39)
- m. Evaluation of processes using ICT (Items 40 and 41)
- n. University support services (Items 42 and 43)
- o. Knowledge of basic concepts, essential computer components, and ICT resources (Items 44, 45, and 46)
- p. Actions carried out by teachers to improve their skills in the use of ICT (Items 47-53)

### Data analysis

The chi-square test was used to verify the hypothesis of whether there is a relationship between the variables to perform the correlational analysis of the data. The alpha coefficient of 0.05 was used. In this sense, the relationships between the sociodemographic variables and the dimensions of digital competencies were analyzed. A certain number of items were applied on a Likert scale to assess the level of digital skills. Likewise, the statistical package SPSS v26 and descriptive statistics were used.

### Ethical issues

Ethical review and approval were waived for this study because it does not involve any risk to the participant's life or health. No substance has been tested on the participants, and they have not been endangered. All procedures performed in studies involving human participants were by the ethical standards of the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

## Results

### Sociodemographic data

The group of participants consisted of 88 men and 34 women. The age groups are shown in Table 1.

Table 1. Demographic characteristics of the sample of professors.

Age	Number of teachers	%
Up to 35 years	6	4.92
From 36 to 45 years	32	26.23
From 46 to 55 years	40	32.79
From 56 to 65 years	34	27.87
From 66 years and older	10	8.20
Teaching experience	Number of teachers	%
Less than 5 years	18	14.75
From 5 to 10 years	30	24.59
From 11 to 20 years	43	35.25
From 21 to 30 years	21	17.21
More than 30 years	10	8.20
Time of dedication	Number of teachers	%
Full-time	24	19.68
Part-time	98	80.32
Hiring category	Number of teachers	%
Hired	113	92.62
Ordinary	9	7.38
<b>Total</b>	<b>122</b>	<b>100%</b>

### Correlational analysis of digital competence with sociodemographic variables

The chi-square test statistic is used to verify the hypothesis of whether there is a relationship or association between the variables, working with an alpha coefficient of 0.005 to perform the correlational analysis of the data. In this sense, the relationships between the sociodemographic variables and the dimensions of digital competencies are analyzed. In this correlation analysis, not all the dimensions of digital competence are presented; the dimension that provides information or where there are findings of greater importance is being considered. Table 1 shows the correlation between sociodemographic variables and knowledge of networking methodological strategies.

When relating the sociodemographic variables with the knowledge and use of networking methodological strategies, we found no relationship between these variables since the significance level in each one is more significant than 0.05 ( $\alpha$ ), as seen in Table 2. And Table 3 shows the correlation between sociodemographic variables and the importance that the teacher gives to the factors when choosing an ICT resource for the classroom.

Table 2. Correlation between sociodemographic variables and knowledge of networking methodological strategies.

Variable	Value	df	Asymptotic significance (bilateral)
<b>Regarding knowledge</b>			
Age	8,264	8	0,408
Sex	0,022	2	0,989
Teaching experience	7,836	8	0,450
<b>Regarding use</b>			
Age	6,368 <sup>a</sup>	8	0,606
Sex	0,177 <sup>a</sup>	2	0,915
Teaching experience	5,084 <sup>a</sup>	8	0,749

Table 3. Correlation between sociodemographic variables and the importance that the teacher gives to the factors when choosing an ICT resource for the classroom.

Variable	Value	df	Asymptotic significance (bilateral)
Age	5,912 <sup>a</sup>	8	0,657
Sex	0,025 <sup>a</sup>	2	0,988
Teaching experience	4,883 <sup>a</sup>	8	0,770

When relating the sociodemographic variables and the importance that the teacher gives to the factors when choosing an ICT resource for the classroom, there is no relationship between these variables since the level of significance in each one is more significant than 0.05 ( $\alpha$ ), as can be seen in Table 3.

Table 4. Correlation between sociodemographic variables and knowledge/use of communication tools and application.

Variable	Value	df	Asymptotic significance (bilateral)
<b>Regarding knowledge</b>			
Age	10,864 <sup>a</sup>	8	0,210
Sex	5,864 <sup>a</sup>	2	0,043
Teaching experience	7,072 <sup>a</sup>	8	0,529
<b>Regarding use</b>			
Age	16,754 <sup>a</sup>	8	0,033
Sex	0,240 <sup>a</sup>	2	0,887
Teaching experience	12,671 <sup>a</sup>	8	0,124

When relating the sociodemographic variables with the knowledge and use of tools and applications related to communication, it was found that there is a relationship between sex and knowledge of the tools since the significance level is lower (Sig. = 0.043 < 0.05 ( $\alpha$ )), as seen in Table 4. Likewise, there is a relationship between age and the use of tools and applications related to communication since the significance level is lower (Sig. = 0.033 < 0.05 ( $\alpha$ )). In addition, in the correlation of the other sociodemographic variables, there is no relationship between these variables since the significance level in each one is more significant than 0.05 ( $\alpha$ ), as seen in Table 4.

Table 5. Correlation between sociodemographic variables and knowledge/use of information tools and applications.

Variable	Value	df	Asymptotic significance (bilateral)
<b>Regarding knowledge</b>			
Age	19,085 <sup>a</sup>	8	0,014
Sex	0,199 <sup>a</sup>	2	0,905
Teaching experience	11,147 <sup>a</sup>	8	0,194
<b>Regarding use</b>			
Age	11,243 <sup>a</sup>	8	0,188
Sex	3,143 <sup>a</sup>	2	0,208
Teaching experience	8,076 <sup>a</sup>	8	0,426

When relating age and knowledge of tools and applications related to information, we found a direct relationship between these variables (Sig. = 0.014 < 0.05 ( $\alpha$ )). There is no relationship between the correlations of the other sociodemographic variables since the significance level in each one is more significant than 0.05 ( $\alpha$ ), as seen in Table 5.

## Discussion

Teachers generally believe their digital competence is high and use it effectively in higher education. A constructive approach should strengthen some deficiencies, but generally, the results are encouraging. While analyzing the teaching of digital competence and its dimensions according to the model proposed by Prendes Espinoza (2010), the model and instrument that have been used in this research, we find certain similarities with the results obtained in earlier research. They include the research conducted by Gutiérrez Porlán (2011), whose study that was conducted with university teachers in Spain, with Carrillo (2014) study conducted with teachers from the Universidad Los Andes in Venezuela and finally Inchaustegui Santoni (2015) whose study was conducted with teachers from UNAPEC in the Dominican Republic.

The analysis showed that the highest scores were found in dimensions d, f and n, such as factors of importance in the choice of ICT resources for the classroom, followed by using the virtual campus and evaluation of processes using ICT. It was found that the lowest scores in dimensions f, k and l, knowledge of the virtual campus, a result that is also contradictory concerning use because a high percentage of teachers uses the tools of the virtual campus. It was also observed that many teachers do not resolve technical issues autonomously and do not participate in research and educational innovation groups. Similarly, it was observed that teachers older than 66 years show less knowledge and use of communication and information tools and applications.

The observation that many teachers struggle with technical issues independently and refrain from participating in research and innovation groups underlines challenges in fostering a self-sufficient and collaborative teaching environment. Additionally, the revelation that educators aged 66 and above exhibit less proficiency in communication and information tools underscores the age-related digital divide, highlighting a crucial area for targeted support and training initiatives. Overall, this analysis provides valuable insights into educators' ICT integration, pinpointing areas for improvement and tailored interventions.

The results show differences between the groups analyzed according to age, sex, and teaching experience. Regarding the age of the participating teachers, it was found in the research of Inchaustegui Santoni (Inchaustegui Santoni, 2015) that the highest percentage of participating teachers is between 40 to 60 years, in Carrillo (2014) between 32 and 42 years, in Gutiérrez Porlán (2011), the teaching staff is younger than 45 years, while in our data obtained, it is from 46 to 55 years. The faculty in Spain and Venezuela are younger than UNAPEC and our study. It can also be observed that younger teachers have a higher perception of their digital competence (Inchaustegui Santoni, 2015) in all dimensions and that there is a significant relationship between age and the use of tools and applications related to communication (Inchaustegui Santoni, 2015). These results also coincide with the study proposed by Esteve (2015).

On the other hand, regarding the sex of the participating teachers, we found in Inchaustegui Santoni (2015), Gutiérrez Porlán (2011), and in our study that the highest percentage of teachers is male, while in Carrillo (2014), the highest percentage is female. With this comparison, we have some evidence that males predominate in university teaching positions. In addition, the highest scores in knowing and using ICT tools correspond to men at a general level, and except for dimensions f, k and l, the score of men was slightly higher, showing minimal and non-significant differences.

Teaching experience plays a vital role in improving teachers' pedagogical practices through continuous learning and exchanging experiences with students. According to Inchaustegui Santoni (2015), the highest percentage of teachers in UNAPEC have experience ranging from 21 to 30 years. On the other hand, Gutiérrez Porlán (2011) found that the teachers' experience is from 11 to 20 years. It is worth noting that the teachers in Spain are younger compared to the other studies analyzed, including ours. However, there is no relationship between teaching experience, knowledge, and methodological networking strategies.

Another characteristic point we have is the time of teaching dedication in Carrillo (2014) and Inchaustegui Santoni (2015); it was found that the highest percentage of teachers is full-time, while in our study, the highest percentage is part-time. Likewise, concerning the distribution by faculty category, we found that in Carrillo (2014), Gutiérrez Porlán (2011), and Inchaustegui Santoni (2015), the highest percentage of teachers are in the category of full professors, compared to our study, where 94% are in the category of contract teachers, i.e., they are not tenured or appointed in the Peruvian context.

Similarly, the significant difference in faculty category distribution, with a majority in the category of contract teachers in our study, suggests a distinct faculty composition compared to the other studies. These variations could impact an institution's overall dynamics of teaching and academic responsibilities. It would be valuable to explore the reasons behind these differences and their potential effects on teaching quality, faculty engagement, and other relevant aspects in the Peruvian context.

Concerning knowledge about essential computer components and basic ICT concepts, we find similar results with Gutiérrez Porlán (2011), Carrillo (2014), and Inchaustegui Santoni (2015), where the highest percentage of the teaching staff values their competence at a superficial level. The only difference is that in our study, for the indicator selection and acquisition of ICT resources, the highest percentage of the participating teaching staff values their competence in the deep level. Likewise, we have very favorable results regarding the use of protection measures for their equipment since the highest percentage of teachers use security measures. In addition, our results are like the ones by Gutiérrez Porlán (2011) and Carrillo (2014), while in Inchaustegui Santoni (2015), it was observed that the highest percentage of teachers do not use security measures.

Regarding knowledge and use of methodological strategies for networking, the results of our study show a similarity with Gutiérrez Porlán (2011) and Carrillo (2014) since the highest percentage of teachers have evaluated their competence that they know and use a little to a lot, while the results of Inchaustegui Santoni (2015) show that the highest percentage of teachers have no knowledge and also do not use the tools.

About the possibilities and limitations of ICT in the teaching-learning process (Prendes Espinoza, 2010), we see that the possibilities or advantages best valued both in Gutiérrez Porlán (2011), Inchaustegui Santoni (2015), Carrillo (2014) and also in our study are the access to information and time flexibility (Prendes Espinoza, 2010). Regarding the most significant limitations, we agree with Gutiérrez Porlán (2011), Carrillo (2014), and Inchaustegui Santoni (2015), where the highest percentage selected user limitations and technical failures. However, among the least valued, we also agree with Gutiérrez Porlán (2011), who reported the flexibility of spaces and publication of information.

About the factors of importance when choosing ICT resources for the classroom, the three studies and ours have coincided in valuing as the most important being the knowledge of the use of the resources, that it solves learning needs, there is ease of access for students and that it is a motivating resource for students. Another aspect where there is a surprising coincidence between the three studies and ours is the dimension knowledge and use of communication and information tools and applications, where they state that they know and use tools such as e-mail, forums, and instant messaging or chat. In the less known ones, Gutiérrez Porlán (2011) and our study coincide in teachers not knowing and using microblogging and virtual worlds, while in Inchaustegui Santoni (2015)'s research, teachers do not know and use file sharing. Likewise, regarding the knowledge and use of information tools, the three studies agree on teachers knowing and using search tools, followed by text editors and visual presentation creators. In contrast, Gutiérrez Porlán (2011) and Inchaustegui Santoni (2015) agree in the teachers not knowing and using live streaming, while in our study, it turned out to be social bookmarking, podcasts, and web page editor. According to Gutiérrez Porlán (2011) and Inchaustegui Santoni (2015), the teaching staff is familiar with and uses the virtual campus extensively.

Regarding the publication of didactic material and scientific production on the network, it was found that in Inchaustegui Santoni (2015)'s paper, teachers sometimes publish, while in Gutiérrez Porlán (2011)'s, the teachers declare to carry out quite a lot of publication actions. About the autonomous ability to solve technical problems on computer equipment, we find significant differences in UNAPEC.

In a recent study, only 9% of respondents claimed to be able to solve technical incidents, while in Spain, the figure was 55%. However, in this study, 78% of respondents stated that they can do so, which is a very encouraging result compared to previous studies. This increase in technical proficiency can be attributed to the current COVID-19 health emergency, which has increased the need for individuals to learn and use ICT tools.

In relation to the use of ICTs for teaching tasks, both Inchaustegui Santoni (2015)'s, Gutiérrez Porlán (2011)'s and our study coincide in using technology tools a lot for electronic administration and management (Prendes Espinoza, 2010).

Regarding the use of virtual tutoring to engage students, there are notable differences between the studies conducted by Gutiérrez Porlán (2011), Inchaustegui Santoni (2015) and our study. Gutiérrez Porlán (2011) reports that 40% of teachers use virtual tutoring frequently, while Inchaustegui Santoni (2015) reports a frequency of 29%. Our study, on the other hand, found that 83% of teachers use virtual tutoring frequently to evaluate students. Regarding using virtual tutoring to evaluate students, we agree with Gutiérrez Porlán (2011) that 80% of teachers use it, as opposed to Inchaustegui Santoni (2015), where only 49% of teachers use it.

Less than 60% of the participants in Gutiérrez Porlán (2011) and Inchaustegui Santoni (2015) studies believed they had effective strategies to encourage student participation. In our study, 86% declare that they have a skill between quite a lot and a lot. However, when analyzing the percentage of those who marked which strategies they apply, we find slight differences with the Spanish teachers since, despite having a lower percentage, they show that they use more strategies compared to UNAPEC and our study. Regarding the assessment of processes using ICT, ICT is used more in the assessment process for the same reason that teachers use a variety of tools that have to do with online questionnaires, followed by creation, understanding and memory, and lastly application.

Regarding the support services by the university, we have agreed with a high percentage of the three authors that the teaching staff declares to know and use the support service of their universities. Regarding teacher training and evaluation of teaching practice with ICT, both Gutiérrez Porlán (2011) and Inchaustegui Santoni (2015) agree to participate often and very often in ICT training actions. However, regarding the frequency in which they impart ICT training, it is observed that more than 50% of the Spanish teachers state that they have never imparted training compared to 29% of UNAPEC and 19% in our study that they imparted training.

Regarding the support services provided by the university, the high percentage of agreement between the three authors is notable, where the teaching staff declares that they know and use the support services of their respective institutions. It is crucial to address these differences to promote comprehensive and practical teacher training, thus guaranteeing a more homogeneous implementation of ICT in teaching.

There are significant differences in the evaluation of their teaching practice since in the study by Gutiérrez Porlán (2011) and Inchaustegui Santoni (2015), less than 50% evaluate their teaching practices with ICT, while in our study, more than 80% of teachers evaluate their practice with ICT. Regarding the actions to improve ICT skills, answers are varied; we observed that UNAPEC teachers outperform teachers in Spain in the use of professional networks, as



well as in disseminating their teaching experience with ICT, while in our study regarding actions to improve ICT skills, the results are disappointing because it was found that 34% of teachers never use digital technology and 54% use sources only sometimes, we also have that 33% who never access digital resource repositories and 50% who access only sometimes.

In a similar situation, 30% never participate in reflection forums, 47% only do so sometimes, and 70% do not participate in professional networks. We also have low participation in research and innovation groups, finding that 19% have never participated and 36% have only participated a few times. Likewise, in the dimension of participation in educational innovation projects, we find that 52% of teachers do not participate.

After the discussion, we can find that it is necessary to continue raising the level of digital competence of university teachers through autonomous learning by the teachers themselves or continuous training by universities (Bozu & Canto, 2009). If we reflect on what is analyzed in this study, we can mention that through digital education, all professionals and, in this case, the teachers can immerse themselves in the current world through technology. As the number of teachers trained in knowing and using ICT tools increases, they can work in collaborative networks, publish research in scientific networks, learn autonomously and independently with the ability to promote creative thinking in their students and reflect on their teaching practice.

University education in Peru and worldwide has experienced a notable digital transformation in recent years, marked by technological advances that have impacted how it is taught and learned. The information provided reveals both advances and challenges in integrating information and communication technologies (ICT) in the university environment. This academic discussion highlights the changes in digital education, the need to develop more significant digital competencies and the crucial role of artificial intelligence (AI) in education and research.

Based on the results of various investigations, it is evident that university teachers in Peru show a favorable perception of their level of digital competence. However, deficiencies that require strengthening from a constructive approach are also highlighted. The findings reveal that teachers show strengths in the choice of technological resources for the classroom, using the virtual campus and evaluating processes through ICT. On the other hand, weaknesses are observed in areas such as knowledge of the virtual campus and autonomous resolution of technical problems. Furthermore, it is worrying that many teachers do not participate in educational research and innovation groups. These results underline the need to promote the development of digital competencies more comprehensively and promote technical autonomy among educators.

The results also reveal significant differences in age, gender and teaching experience. Regarding age, it stands out that younger teachers tend to perceive their digital competence more positively. This correlation between youth and digital competence is consistent with previous studies. Furthermore,

a male predominance is identified in university teaching positions, which suggests possible challenges related to gender equity in this area. Although teaching experience is vital to improving pedagogical practices, it does not present a transparent relationship with digital competence. This finding highlights the importance of addressing digital divides through continuing education programs, regardless of accumulated teaching experience.

It is crucial to analyze teachers' knowledge and use of specific digital tools in digital education. Choosing classroom resources, publishing teaching materials online, and participating in university support services are areas in which teachers show strengths. However, challenges persist in implementing methodological strategies for interconnection (networking) and the stimulation of student participation.

The research reveals that the perceived limitations when using ICT in the teaching-learning process focus on the limitations of the users and technical failures. However, the most valued possibilities include access to information and flexibility over time, highlighting the importance of integrating digital education to take advantage of these advantages.

## Conclusion

The results obtained in this research reveal a worrying panorama regarding the digital competence of the teachers surveyed and analyzed. A general level that ranges between low and medium, with some indicators standing out at the high level, is evident. These findings pose a significant challenge since today's society is immersed in a reality where education has been strongly impacted, demanding the presence of teachers with solid digital skills to face the various educational demands. It is crucial to highlight that the dimensions related to the knowledge and use of methodological strategies for interconnection, as well as the mastery of communication and information tools and applications, the publication of didactic material and scientific production on the Internet, together with the ability to solve technical problems in their technological equipment autonomously, reflect worryingly low results. These aspects are fundamental in today's educational environment, where technology plays a crucial role in pedagogical effectiveness.

Additionally, disparities were identified in teachers' digital competence based on gender and age, with the relationship between gender and knowledge of interconnection methodological strategies being the most significant. Likewise, a correlation was evident between age and the use of tools and applications related to communication and between age and knowledge of tools and applications linked to communication. Younger teachers perceive themselves as more competent in using digital tools, while older teachers report using them infrequently or not at all.

The authors hope this research will not only shed light on the current situation but also provide practical proposals to improve the quality of the teaching staff, the educational institution and, ultimately, student learning. Furthermore, in

line with the principles of educational design, it is intended that this study has an applicable value in different contexts and contributes to the knowledge of society as a whole, transcending the limits of the specific institution in which it was carried out.

## References

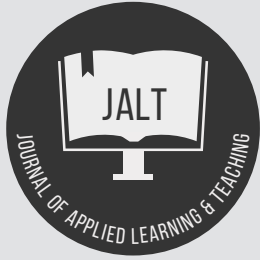
- Ageyi, D. D. (2021). Integrating ICT into schools in Sub-Saharan Africa: From teachers' capacity building to classroom implementation. *Education and Information Technologies*, 26(1), 125-144. <https://doi.org/10.1007/s10639-020-10253-w>
- Alt, D. (2018). Science teachers' conceptions of teaching and learning, ICT efficacy, ICT professional development and ICT practices enacted in their classrooms. *Teaching and Teacher Education*, 73, 141-150. <https://doi.org/10.1016/j.tate.2018.03.020>
- Amhag, L., Hellström, L., & Stigmar, M. (2019). Teacher educators' use of digital tools and needs for digital competence in higher education. *Journal of Digital Learning in Teacher Education*, 35(4), 203-220. <https://doi.org/10.1080/21532974.2019.1646169>
- Antonietti, C., Cattaneo, A., & Amenduni, F. (2022). Can teachers' digital competence influence technology acceptance in vocational education? *Computers in Human Behavior*, 132, 107266. <https://doi.org/10.1016/j.chb.2022.107266>
- Atmacasoy, A., & Aksu, M. (2018). Blended learning at pre-service teacher education in Turkey: A systematic review. *Education and Information Technologies*, 23(6), 2399-2422. <https://doi.org/10.1007/s10639-018-9723-5>
- Aylett, R., & Gregory, K. (1997). *Criteria for teaching competence and teaching excellence in higher education*. Lalmer Press.
- Böttcher-Oschmann, F., Groß Ophoff, J., & Thiel, F. (2021). Preparing teacher training students for evidence-based practice promoting students' research competencies in research-learning projects. *Frontiers in Education*, 6, 642107. <https://doi.org/10.3389/feduc.2021.642107>
- Bozu, Z., & Canto, P. J. (2009). University professors in the knowledge society: Professional teaching competencies. *Revista de formación e innovación educativa universitaria*, 2(2), 87-97.
- Bunk, G. (1994). The transfer of skills in professional training and further education in the FRG. *Revista europea de formación profesional*, (1), 8-14.
- Bustos López, H. G., & Gómez Zermeño, M. G. (2018). Digital competence in high school teachers as a means for educational innovation. CPU-e. *Revista de Investigación Educativa*, (26), 66-86.
- Cabero-Almenara, J., Gutiérrez-Castillo, J.-J., Palacios-Rodríguez, A., & Barroso-Osuna, J. (2020). Development of the teacher digital competence validation of digcompedu check-in questionnaire in the university context of Andalusia (Spain). *Sustainability*, 12(15), 6094. <https://doi.org/10.3390/su12156094>
- Carrillo, D. (2014). *ICT competencies of teachers for teaching through virtual environments in higher education. The case of Universidad de los Andes-Venezuela: Evaluation and design of a training plan*. Universitat Rovira I Virgili, Catalunya. <https://www.tdx.cat/handle/10803/285330#page=1>
- Castillo-Martínez, I. M., & Ramírez-Montoya, M. S. (2021). Research competencies to develop academic reading and writing: A systematic literature review. *Frontiers in Education*, 5, 576961. <https://doi.org/10.3389/feduc.2020.576961>
- Chávez-Miyauchi, T.-E., Benítez-Rico, A., Alcántara-Flores, M., Vergara-Castañeda, A., & Ogando-Justo, A.-B. (2021). Personal motivation and learning self-management in students, as result of the transition to online courses during COVID-19 pandemic. *Nova scientia*, 13(SPE), 0-0.
- Chuang, S. (2021). The applications of constructivist learning theory and social learning theory on adult continuous development. *Performance Improvement*, 60(3), 6-14. <https://doi.org/10.1002/pfi.21963>
- De la Calle, A. M., Pacheco-Costa, A., Gómez-Ruiz, M. Á., & Guzmán-Simón, F. (2021). Understanding teacher digital competence in the framework of social sustainability: A systematic review. *Sustainability*, 13(23), 13283. <https://doi.org/10.3390/su132313283>
- De la Fuente, C. I., Guadagnin, E. C., Kunzler, M. R., & Carpes, F. P. (2021). Programming course for health science as a strategy to engage students during the coronavirus pandemic. *Advances in Physiology Education*, 45(1), 53-58. <https://doi.org/10.1152/advan.00183.2020>
- DeVries, R. (2000). Vygotsky, Piaget, and education: A reciprocal assimilation of theories and educational practices. *New ideas in Psychology*, 18(2-3), 187-213. [https://doi.org/10.1016/S0732-118X\(00\)00008-8](https://doi.org/10.1016/S0732-118X(00)00008-8)
- Esteve, F. (2015). *Digital teaching competence. Analysis of the self-perception and performance evaluation of university education students through a 3D environment*. Universitat Rovira I Virgili. <https://www.tdx.cat/bitstream/handle/10803/291441/tesis.pdf>
- European Commission. (2010). *Common European principles for teacher competencies and qualifications. Directorate-general for education and culture*. <http://www.pef.uni-lj.si/bologna/dokumenti/eu-common-principles.pdf>
- European Commission. (2012). *Supporting the teaching professions for better learning outcomes. accompanying the document communication from the commission "rethinking education: Investing in skills for better socio-economic outcomes"*. <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2012:0374:FIN:EN:PDF>

- Falloon, G. (2020). From digital literacy to digital competence: The teacher digital competency (TDC) framework. *Educational Technology Research and Development*, 68(5), 2449-2472. <https://doi.org/10.1007/s11423-020-09767-4>
- Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2020). Digital competences for teacher professional development. Systematic review. *European Journal of Teacher Education*, 1-19. <https://doi.org/10.1080/02619768.2020.1827389>
- Feyzi Behnagh, R., & Yasrebi, S. (2020). An examination of constructivist educational technologies: Key affordances and conditions. *British Journal of Educational Technology*, 51(6), 1907-1919. <https://doi.org/10.1111/bjet.13036>
- Finnish Institute for Educational Research. (2010). *Education and training 2010: Three studies to support school policy*. <http://hdl.voced.edu.au/10707/280890>
- Gagné, R. M., & Briggs, L. J. (1974). *Planning teaching*. Ciudad de México: Editorial Trillas.
- Galdeano Bienzobas, C., & Valiente Barderas, A. (2010). Professional competencies. *Educación química*, 21(1), 28-32.
- Gardner, H. (1982). Giftedness: Speculations from a biological perspective. *New Directions for Child and Adolescent Development*, 1982(17), 47-60. <https://doi.org/10.1002/cd.23219821706>
- Garzón Artacho, E., Martínez, T. S., Ortega Martín, J. L., Marín Marín, J. A., & Gómez García, G. (2020). Teacher training in lifelong learning—the importance of digital competence in the encouragement of teaching innovation. *Sustainability*, 12(7), 2852. <https://doi.org/10.3390/su12072852>
- Gobierno del Perú. (2020). *Legislative decree 1495*. <https://busquedas.elperuano.pe/normaslegales/decreto-legislativo-que-establece-disposiciones-para-garantizar-decreto-legislativo-n-1495-1866211-2/>
- González-Gutierrez, V., Alvarez-Risco, A., Estrada-Merino, A., Anderson-Seminario, M. d. I. M., Mlodzianowska, S., Del-Aguila-Arcenales, S., & Yáñez, J. A. (2022). Multitasking behavior and perceptions of academic performance in university business students in Mexico during the COVID-19 pandemic. *International Journal of Mental Health Promotion*, 24(4), 565-581. <https://doi.org/10.32604/ijmh.2022.021176>
- GutiérrezPorlán, I. (2011). *Competencies of university professors in relation to the use of information and communication technologies: Analysis of the situation in Spain and proposal of a training model*. Universidad Universitat Rovira I Virgill. <https://www.tdx.cat/handle/10803/52835#page=4>
- Habermas, J. (1984). *Theory of communicative action, Volume one: Reason and the rationalization of society* (T. A. McCarthy, Trans.). Beacon Press. (Original work published 1981).
- Habermas, J. (1987). *Theory of communicative action, Volume two: Lifeworld and system: A critique of functionalist reason* (T. A. McCarthy, Trans.). Beacon Press. (Original work published 1981).
- Hamadi, M., El-Den, J., Azam, S., & Sriratanaviriyakul, N. (2022). Integrating social media as cooperative learning tool in higher education classrooms: An empirical study. *Journal of King Saud University - Computer and Information Sciences*, 34(6, Part B), 3722-3731. <https://doi.org/10.1016/j.jksuci.2020.12.007>
- Howard, S. K., Tondeur, J., Ma, J., & Yang, J. (2021). What to teach? Strategies for developing digital competency in preservice teacher training. *Computers & Education*, 165, 104149. <https://doi.org/10.1016/j.compedu.2021.104149>
- Inchaustegui Santoni, A. (2015). *Perception of the ICT competencies of Dominican university professors*. UNAPEC Case Universidad de Murcia. <https://digitum.um.es/digitum/handle/10201/47542>
- Ivenicki, A. (2021). COVID-19 and multicultural education in Brazil. *Perspectives in Education*, 39(1), 231-241. <https://doi.org/10.18820/2519593X/pie.v39.i1.14>
- Martinez, E. G., Padrón, R. R., & Villalba, P. J. (2022). The students' point of view on the teaching of anatomy at the universidad del norte, Colombia, amid the Covid-19 pandemic. *International Journal of Morphology*, 40(1). <https://doi.org/10.4067/S0717-95022022000100046>
- McClelland, D. C. (1973). Testing for competence rather than for "intelligence". *American Psychologist*, 28(1), 1-14. <https://doi.org/10.1037/h0034092>
- MINEDU. (2020a). *Guidelines for the continuity of university higher education services, within the framework of the national health emergency established by Supreme Decree N° 008-2020-SA*. [https://cdn.www.gob.pe/uploads/document/file/574842/RVM\\_N\\_085-2020-MINEDU.pdf](https://cdn.www.gob.pe/uploads/document/file/574842/RVM_N_085-2020-MINEDU.pdf)
- MINEDU. (2020b). *Technical standard provisions for the prevention, care and monitoring of coronavirus (COVID-19) in universities nationwide*. [https://cdn.www.gob.pe/uploads/document/file/565533/RVM\\_N\\_081-2020-MINEDU.PDF](https://cdn.www.gob.pe/uploads/document/file/565533/RVM_N_081-2020-MINEDU.PDF)
- MINEDU. (2021). *Implementation of the continuing education and training program during the year 2021*. <https://cdn.www.gob.pe/uploads/document/file/1792114/RVM%20N%C2%B020109-2021-MINEDU.pdf.pdf>
- MINEDU. (2023). *The university in figures*. <https://repositorio.minedu.gob.pe/bitstream/handle/20.500.12799/9077/La%20Universidad%20en%20Cifras.pdf?sequence=1&isAllowed=y>
- Miró Vera, Y. A. (2019). Diagnostic of self-assessment of students of the design module from the projects area of FUNIBER about IPMA 4.0 competences certification. *Project, Design and Management*, 1(1), 161. <https://doi.org/10.35992/mlspdm.v1i1.161>
- Mulder, M., Weigel, T., & Collings, K. (2008). The concept of competence in the development of vocational education and training in some EU Member States: A critical analysis.

- Profesorado. *Revista de Currículum y Formación de profesorado*, 12(3), 1-25.
- Muñoz, F. I. (2006). The teaching profession in globalization and the knowledge society. In *[La formación del profesorado y la mejora de la educación: Políticas y prácticas [Teacher education and the improvement of education: Policies and practices]* (pp. 231-244). [http://www.ub.edu/obipd/docs/la\\_profesion\\_docente\\_en\\_la\\_globalizacion\\_y\\_la\\_sociedad\\_del\\_conocimiento\\_imbernon\\_f.pdf](http://www.ub.edu/obipd/docs/la_profesion_docente_en_la_globalizacion_y_la_sociedad_del_conocimiento_imbernon_f.pdf)
- Naranjo, Á., Celis, L., & Blandón, O. (2017). Professional teaching competencies: A review of the meaning from different perspectives. *Revista educación y pensamiento*, 24, 6-17.
- OECD. (2009). *Teachers matter: Attracting, developing and retaining effective teachers*. [https://read.oecd-ilibrary.org/education/politica-de-educacion-y-formacion-los-docentes-son-importantes\\_9789264046276-es#page4](https://read.oecd-ilibrary.org/education/politica-de-educacion-y-formacion-los-docentes-son-importantes_9789264046276-es#page4)
- Palan, R. (2003). *Competency management: A practitioner's guide, specialist management resource*. Perpustakaan Negara Malaysia Cataloguing.
- Parlamento Europeo. (2007). *Key competences for lifelong learning - A European framework is the annex to a recommendation of the European Parliament and of the council of 18 December 2006 on key competences for lifelong learning published in the official journal of the European Union L 394 of 30 December 2006*. <https://www.educacionyfp.gob.es/dctm/ministerio/educacion/mecu/movilidad-europa/competenciasclave.pdf?documentId=0901e72b80685fb1>
- Pekkala, K., & van Zoonen, W. (2022). Work-related social media use: The mediating role of social media communication self-efficacy. *European Management Journal*, 40(1), 67-76. <https://doi.org/10.1016/j.emj.2021.03.004>
- Perrenoud, P. (2004). *Ten new skills for teaching: An invitation to travel* (Vol. 196). Graó.
- Perrotta, D. (2021). Universities and Covid-19 in Argentina: From community engagement to regulation. *Studies in Higher Education*, 46(1), 30-43. <https://doi.org/10.1080/03075079.2020.1859679>
- Pöntinen, S., & Rätty-Záborszky, S. (2020). Pedagogical aspects to support students' evolving digital competence at school. *European Early Childhood Education Research Journal*, 28(2), 182-196. <https://doi.org/10.1080/1350293X.2020.1735736>
- Pozo-Sánchez, S., López-Belmonte, J., Rodríguez-García, A.-M., & López-Núñez, J.-A. (2020). Teachers' digital competence in using and analytically managing information in flipped learning. *Culture and Education*, 32(2), 213-241. <https://doi.org/10.1080/11356405.2020.1741876>
- Prata-Linhares, M. M., Cardoso, T. d. S. G., Lopes-Jr, D. S., & Zukowsky-Tavares, C. (2020). Social distancing effects on the teaching systems and teacher education programmes in Brazil: Reinventing without distorting teaching. *Journal of Education for Teaching*, 46(4), 554-564. <https://doi.org/10.1080/02607476.2020.1800406>
- Prendes Espinoza, M. P. (2010). *ICT skills for teaching in Spanish public universities: Indicators and proposals for the definition of good practices: Study and analysis program. Report of project EA2009-0133 of the secretary of state for universities and research*. <http://www.um.es/competenciastic>
- Qaddumi, H., Bartram, B., & Qashmar, A. L. (2021). Evaluating the impact of ICT on teaching and learning: A study of Palestinian students' and teachers' perceptions. *Education and Information Technologies*, 26(2), 1865-1876. <https://doi.org/10.1007/s10639-020-10339-5>
- Ricaurte, M., Ordóñez, P. E., Navas-Cárdenas, C., Meneses, M. A., Tafur, J. P., & Viloría, A. (2022). Industrial processes online teaching: A good practice for undergraduate engineering students in times of COVID-19. *Sustainability*, 14(8), 4776. <https://doi.org/10.3390/su14084776>
- Rojas-Osorio, M., & Alvarez-Risco, A. (2019). Intention to use smartphones among Peruvian university students. *International Association of Online Engineering*, 13(3), 40-52. <https://doi.org/10.3991/ijim.v13i03.9356>
- Rossi Cordero, A. S., & Barajas Frutos, M. (2018). Digital competence and educational innovation: Challenges and opportunities. *Profesorado - Revista de Currículum y formación de profesorado*, 22(3), 317-339.
- Rybczewska, M., & Sparks, L. (2022). Ageing consumers and e-commerce activities. *Ageing and Society*, 42(8), 1879-1898. <https://doi.org/10.1017/S0144686X20001932>
- Santos, L. M., Grisales, D., & Suero Rico, J. (2021). Perception and technological accessibility of university students in the southwest of the Dominican Republic during Covid-19. *Revista Internacional de Educación para la Justicia Social*, 10(1), 145-165. <https://doi.org/10.15366/riejs2021.10.1.009>
- Sepulveda-Escobar, P., & Morrison, A. (2020). Online teaching placement during the COVID-19 pandemic in Chile: Challenges and opportunities. *European Journal of Teacher Education*, 43(4), 587-607. <https://doi.org/10.1080/02619768.2020.1820981>
- Skantz-Åberg, E., Lantz-Andersson, A., Lundin, M., & Williams, P. (2022). Teachers' professional digital competence: An overview of conceptualisations in the literature. *Cogent Education*, 9(1), 2063224. <https://doi.org/10.1080/2331186X.2022.2063224>
- Skinner, B. F. (1981). *Verbal behavior*. Trillas México.
- Soegoto, E. S., & Eliana, E. (2018). E-commerce and business social media today. *IOP Conference Series: Materials Science and Engineering*, 407, 012034. <https://doi.org/10.1088/1757-899x/407/1/012034>
- Suárez-Rodríguez, J., Almerich, G., Orellana, N., & Díaz-García, I. (2018). A basic model of integration of ICT by teachers: Competence and use. *Educational Technology*

- Research and Development*, 66(5), 1165-1187. <https://doi.org/10.1007/s11423-018-9591-0>
- SUNEDU. (2023). *Universities*. <https://www.sunedu.gob.pe/lista-universidades/>
- Trujillo-Segoviano, J. (2014). The competency-based approach and the improvement of education. *Ra Ximhai*, 10(5), 307-322.
- UNESCO. (1998). *Higher education in the twenty-first century: Vision and action*. World Conference on Higher Education. París. <http://unesdoc.unesco.org/images/0011/001163/116345e.pdf>
- UNESCO. (2009). *ICT competency framework for teachers*. <https://es.unesco.org/themes/tic-educacion/marco-competencias-docentes>
- UNESCO. (2015). *Joint ILO/UNESCO committee of experts on the application of the recommendations concerning teaching personnel*. UNESCO Conference. <https://unesdoc.unesco.org/ark:/48223/pf0000235313>
- Wisevoter. (2023). *Internet speed by country*. <https://wisevoter.com/country-rankings/internet-speed-by-country/#peru>
- Yot Dominguez, C., & Marcelo, C. (2005). *Study on professional competencies for e-learning* (G. d. i. I.D.E.A & E. Universidad de Sevilla, Eds.). Junta de Andalucía. Consejería de empleo. Dirección General de Formación para el Empleo. [https://idus.us.es/bitstream/handle/11441/41183/Competencias\\_profesionales\\_para\\_el\\_desempe%C3%B1o\\_en\\_eLearning.pdf?sequence=1&isAllowed=y](https://idus.us.es/bitstream/handle/11441/41183/Competencias_profesionales_para_el_desempe%C3%B1o_en_eLearning.pdf?sequence=1&isAllowed=y)
- Zabalza, M. Á., & Beraza, M. Á.Z. (2003). *Teaching competencies of university professors: Quality and professional development* (Vol. 4). Narcea Ediciones.
- Zúñiga Rodríguez, M., & Cáceres Mesa, M. L. (2021). The school sense in the face of covid-19. The perception of students from public universities in Hidalgo. *Conrado*, 17(78), 46-53.

Copyright: © 2024. Mercedes Rojas-Osorio, Shyla Del-Aguila-Arcetales and Aldo Alvarez-Risco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## 'Just get them over the line': Neoliberalism and the execution of 'excellence'

Martin Benedict  
Andrew<sup>A</sup>

A

*Associate Professor of Work-Based Learning, Otago Polytechnic, Dunedin, Aotearoa/New Zealand*

---

### Keywords

Critical evocative autoethnography;  
excellence;  
managerialism;  
narrative enquiry;  
neoliberalism;  
professional practice;  
timeliness;  
zombilingo.

---

### Abstract

Higher Education institutions, at least in nominal 'Western' contexts, oversell a dressed-up version of 'excellence' to draw in students, but they have more firmly in their sight a vision of the commercial bottom line. This research study, firmly grounded in the author's experience of postgraduate education, posits that the marketised, neoliberal conception of 'excellence' both covers a hidden truth that these institutions are content just to get learners 'over the line' and hides a more authentic, bottom-up conception of 'excellence' which appears when the voices of learners and educators are heard above the managerialist chatter and when teaching well is considered. The paper presents its review of the field as a critically evocative autoethnography, with the author positioned as a witness to the lexical slaughter of 'excellence', amongst other terms suborned by the neoliberalist academy.

---

### Correspondence

[martin.andrew@op.ac.nz](mailto:martin.andrew@op.ac.nz)<sup>A</sup>

Within its chorus of voices, it introduces the reflective critical incidents of three postgraduate supervisors or mentors telling of their realisations that their institutions are more interested in getting students 'over the line' in a timely fashion than in facilitating opportunities for authentic excellence. Methodologically, then, the study presents three narratives as evidence in a narrative enquiry embedded in the broader autoethnography, as is often the case in professional practice research. The paper is positioned on the cusp of COVID-19's aftermath and suggests that higher education organisations are on track to miss the opportunities to begin dismantling neoliberal thought that the pandemic afforded. Instead, they condemn themselves to 'mediocracy' – rule by the mediocre churning out mediocrity under the guise of a shopfront of gaudy but vacuous 'excellence'.

---

### Article Info

Received 22 February 2024  
Received in revised form 1 March 2024  
Accepted 4 March 2024  
Available online 18 March 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.25>

## Hard times for higher education

Educators in higher education (HE) contexts in the 2020s are forced to toe (not tow) an ever more unpalatable line. This is the line where the learner and the educator's lived experience of 'excellence' meets what a neoliberalist would construe as *the bottom line*. In this context, the 'bottom line' is where managers who are forced to embody the ideology of 'just get them over the line' lurk, knowing that a sea of C-pluses is all it will take to ensure they meet their own bottom lines in performance, department, faculty reviews and (*bingo!*) Key Performance Indicators (KPIs). The bottom line, I'd say, is where mediocrity dwells. Catalysing my exit from academia in Australia in 2016 because of my perception of the *mediocracy* (*sic*) of the sector, an article in *The Age* was headed 'Bottom line always at the top for neoliberalism' (*The Age*, 2016). Cajoled to get students 'over the line' in deference of the bell curve, I realised that the line was actually the bottom line, and how low a bar can you have?

This paper finds its gap and its line of enquiry from my professional practice as a postgraduate educator in Australia and New Zealand. My enquiry is supported by geopolitical thinkers such as Giroux (2009, 2014, 2017, 2019), who viewed neoliberalism as a wartime occupier of higher education; Barnett (2013), who (re)imagined the university; Collini (2012), who asked what universities are for and Fleming (2021) who, like myself (Andrew, 2022), has seen the dark side of the underbelly of the modern (read neoliberal) university. This paper, largely a critical content review, takes critical evocative autoethnography as its method of presentation, incorporating other supervisors' narrative enquiries. I bring in these short mentor narratives, vignettes really, into the evidence space to support the evocative autoethnography and provide some instances of the execution of excellence in practice. I suggest that mediocrity, and not excellence, is the new bottom line of the neoliberal university, despite the expectations of 'excellence' HE institutions sell in Western nations (Moore et al., 2017).

Such institutions exploit, I think, dishonesty and a perception gap between measured 'excellence' and what seems to potential learners to be a more authentic educative excellence. This gap was only amplified by HE institutions' scrambling back after the COVID-19 lockdowns of the early 2020s, using the neoliberalist boost to train compliant workers (Waller & Wrenn, 2023). Measured excellence is seen as a *smoke screen* for driving neoliberalist sectoral change (Cui et al., 2021). I posit that a more bottom-up but robust consideration of what facilitation, mentoring or teaching well could look like in higher educational contexts in the 2020s (see Brookfield et al., 2023) counterbalances to some extent neoliberal biases inherent in the hegemonic concept of 'excellence' and its adjunct, 'quality assurance'.

## The unreality of excellence

Excellence in education is not a commodity that can be bought and sold and made available to all at will. Instruments evaluating excellence in any context will always be vested and never neutral, as with the UK's Teaching Excellence

Framework (TEF) in 2016. We may see agenda-pushing or non-neutrality in Aotearoa/New Zealand's National Centre for Tertiary Teaching Excellence, Australia's ERA (Excellence in Research Australia) or any Vice Chancellor's excellence award in any given year, the annual academic Oscars. Then there are the transformative Academic Excellence Initiatives (AEIs) in elite universities in nine non-Anglophone countries detailed in *Academic star wars* (Yudkevich et al., 2023). The European Association for Quality Assurance in Higher Education (EAQHE; Brusoni et al., 2004), offering a varied discussion of excellence, emphasised the vested nature of 'excellence': "influence of expectations and priorities from different groups can be easily detected when statements about excellence in higher education are drafted" (p. 20). It also admitted that excellence must have quantitative and qualitative parameters and objective and subjective indicators. There is a strong sense that standards-based excellence still dominates the discourse despite the lexical concession. Their claim that excellence is "a reality" and that there is a different definition for each reality (p. 21) brings metaphysical concern for what *reality* is back to the table. One reality, like one size, one 'excellence', can never fit all.

Micromanaged evaluation aside, excellence in education can be seen as the result of the experience, passion and experience of educators and their own ability to both portray this and pass it on (Wood & Su, 2017). This is a more authentic excellence because it comes from a deep-seated place of lived educator experience, motivated by a desire to transform lives and communities (Carpenter & Ker, 2017). The voices of these educators are, as Cui et al. (2021) have shown, too often ignored in top-down neoliberalist conceptualisations of excellence inscribed in Higher Education, such as the TEF. Canning (2019) labelled the policy of the TEF a *hyperreal simulacrum* because it does not reflect teaching excellence but a value-for-money, bottom-line agenda related to graduate employability, another contested subject. In other words, it is inauthentic excellence. Tinto (2017) suggested that viewing excellence through the eyes of learners might improve retention rates.

Authentic educative excellence is experienced, situated, socially enacted and co-constructed in practice. It moves beyond teaching excellence in that the learner's experience determines the authenticity. Their interfaces are the educator, the educator's team, the curriculum, including the assessment process and its transparency and perceived fairness, and any media by which communication about enrolment, programme, assessments and results travel, such as learning management systems and administrative emails. Skelton (2005) notably proposed four dimensions of teaching excellence: traditional, as in sage-on-stage inculcation; performative, where learners are seen as returns on investment and defined by standardised or market-driven quantitative criteria; psychologised, emphasising theory of learning and teaching; and, finally, critical, where understandings of teaching excellence link it to the emancipatory powers of freedom, justice and self-empowerment.

Beyond the conventional and the measurable, teaching excellence is largely the result of educators' enthusiasm/inspiration of those they co-negotiate educational goals with:

the learners. It is often the educator's passion which learners remember long term – call it *Goodbye Mr Chips* syndrome if you will (after Hilton, 1933). What the learners themselves bring to the excellence table matters, too. For instance, such groups as self-motivated and adult learners may embody genetic dispositions and learned traits that may facilitate the co-construction of excellence in their application, work and performance. In conscious danger of bankrupting Bourdieu's idea of cultural capital (1986), I suggest they may have the *habitus* ("socially constituted cognitive capacity"; p. 27) of an excellent learner, just as educators may embody a *habitus* for excellence as educators.

"Teaching excellence" is obviously a contested term (Skelton, 2004), inflected by a regulatory voice invested in setting fees and marketing plans pivoting on "student education" and "student choices" as much as being a mechanism for the imperatives of teaching excellence and the quality assurance that implies (Gunn, 2018; Moore et al., 2017; Owens et al., 2021). A key issue in top-down configurations of excellence in education is its minimalisation of collaboration with the teacher/mentor voice. Wood and Su (2017) were thus motivated to create an empirical yet nuanced study of excellence based on 16 teacher voices with a focus on practice, a methodological principle mirrored by Goode (2023). She writes:

If organisational leaders were asked how they measure the impact of teaching, they would probably refer to key performance indicators, such as retention rates, completion levels, and destination surveys... However, I would argue that, while those measures are certainly important, if learners were asked the same question, these elements would not come into play (p. 1).

Wood and Su (2017) demonstrate that cross- and trans-institutional shared collaboration between educators and learners epitomises excellence in teaching in educators' views. Further, Wood and Su (2017) and Goode (2023) show, amongst other things, that excellent educators are given opportunities to nurture their excellence through research, so that the research: practice nexus becomes a mutually fruitful space. What makes research excellent is subject to scrutiny, too. Tracey (2010), for instance, offered the eight 'big-tent' criteria for judging excellence in qualitative research: worthy topic; rich rigour; sincerity; credibility; resonance; significant contribution; ethical; and meaningful coherence. If collaboration and applied research are supported, the institutions where they occur might more successfully become facilitators and mediators of learning. Thus, they are more likely to afford the features of authentic excellence that support teaching and learning.

It is worth adding that a current report in *The Chronicle of Higher Education* (Elias & O'Leary, 2023) emphasises that few people rate the performance of higher education providers as 'excellent' (only 9%) with a majority plumping for 'good' (41%) or 'very good' (31%) and 19% rating negatively. In an otherwise colonised critique of excellence, Grifoll (2016), a key architect of EAQAHE in 2014, ultimately understands excellence "as a link between innovation and the aim of moving up to better societies" (p. 96). This fits the view

that a more authentic excellence lies in what is fresh and distinctive about practice and in what may be genuinely transformative.

## The vacuousness of excellence

The discursive peddling of excellence as part of a commodity package goes back years and its ghost haunts us post-COVID. Thirty-six years ago, Timar and Kirp (1988) argued that centralised over-regulation kneecapped true innovation and change in the national quest for educational excellence in the United States. In this quest, rules, they write, "cannot compel teachers to be more caring" and cannot "require... administrators be fair and just" (p. 39). Thirty-five years ago, Woudstra and Powell (1989) described those responsible for the appearance of excellence in Higher Education in neoliberalist terms of *competitive advantage*. They wrote that when the services of highly competent academics and tutors, registry staff, student advisors, and counsellors combine, a unique bond is formed between the university and its learners. This unique bond, they argue, becomes a differentiating *competitive advantage* when the institution subscribes to a vision of quality, support, service, and, of course, consequently, 'excellence'.

Collini (2012) identifies the 'vacuity' associated with 'excellence' used in such contexts of raising standards and improving quality. Pointing to the impossibility of the ever-more exponential growth curve, he writes: "Vacuity is... rendered more vacuous still by the requirement that the 'excellent' must become 'yet more excellent'" (p. 109). Trying to quantify the non-scalable concept of excellence is logically fallacious. Moore et al. (2017) saw quantified excellence as fetishised in educational sloganeering: "Excellence R Us." Clegg (2007) captures the oxymoronic nature of quantifying and scaling excellence:

Excellence has become ubiquitous as a popular slogan, indeed the oxymoron 'excellence comes as standard' has thrown off its ironic resonance and is now routinely used to promote an astonishing variety of goods (p. 91).

There will inevitably be factors beyond the agency of the stakeholders who play a role in experienced authentic excellence, as epitomised by COVID-19 and its lockdowns, but how these parties perform and offer support during the unforeseen is also a measure of their commitment to excellence. Key themes about crisis leadership during COVID are the need for clear communication, the quality of compassion (Tan, 2022) and creating spaces where resilience may flourish (Andrew et al., 2020; Balasubramanian & Fernandes, 2022). It is cooperation, not competition, that breeds sustainable resilience (van Staveren, 2023).

## The lexical slaughter of excellence

Henri Giroux, the most prolific opponent of the neoliberalised university, wrote scathingly of the commodification of excellence as part of a hard-sell package: "the appeal to excellence... functions like a corporate logo, hyping



efficiency while denuding critical thought and scholarship of any intellectual, civic, and political substance" (2009, p. 673). Barnett (2013) regarded exercises of university (re) branding and as forms of imaginative, ideological and ethical constraint, noting they are marked by the hijacking of terms such as 'flexible delivery', 'lifelong learning' and a generic branded form of 'excellence'. These, Katz (2015) argues, are "equivocations, neologisms and business-oriented euphemisms" (p. 557). While marketisation might have led to practical improvement, Katz argues, it instead fumbles into imaginative paucity. Katz (2015) also notes the neoliberal university found itself constrained by 'Darwinian' competition towards what Barnett had called 'global excellence' (2013, p. 58). This is the lexical slaughter Watson (2003) identified.

Katz (2015) is one of many authors whose thesis is that neoliberal, marketised and corporate language usurps and distorts the primary purpose of higher education to empower future generations with applied problem-solving strategies, made rigorous via reflective and critical thinking skills. Katz cites Swinburne University in Melbourne, an organisation I left in the mid-2010s because of its mission slip during a discussion of 'behaviours', ways members of the body politic were expected to behave and hence be judged. When a colleague suggested *compassion* had a role to play in our interactions with learners and colleagues, he was severely laughed down. I stood with Tan (2022): "To be a compassionate, inclusive, and mindful educator, it is important for us to be authentic in the positioning of our teaching values and interactions" (p. 157). This dissing of compassion was, for me what led to the critical moment, the realisation of *mediocracy*, I mentioned at the start. In Swinburne's marketing, Katz (2015), too, saw instances of shallow corporate-speak:

Instances of Zombilingo may be found in the 'Swinburne Behaviours', where staff assessment criteria encourage employees to 'support, empower and encourage others to achieve excellence'; perhaps because as previously mentioned, 'together we can make a difference' (p. 56).

What repels me here is the usurping of the social mission of higher education, *making a difference*, to the marketised discourse of excellence. The pretence for collegiality and community is buried in a discourse no longer of shared mission or values but individual behaviours. Lorenz (2012) warned us about the slippage of language in new public management, the handmaiden of neoliberalism: "New public management... parasitises the everyday meanings of (its) concepts... and simultaneously perverts all their original meanings" (Lorenz, 2012, p. 600). The "bullshitter", Lorenz (2012) tells us, "is only interested in effects and does not necessarily believe in what [they] state [themselves]" (p. 560). Analysing documentation in a Dutch context, he writes that a particular variety of this bullshit-lingo is "excellencespeak" (Lorenz, 2012, p. 626).

When Katz (2015) reminded me of organisational behaviours, the memory brought to mind the Pavlovian zombification of higher education, with automaton inhabitants barred from autonomy, free speech and agency and forced into

corporatised models of being, speaking and behaving, often called being a 'team player', like the human outputs Waller and Wrenn (2023) identify. Team players are "robust, resilient, responsive, flexible, innovative, and adaptable" (Gillies, 2011, p. 210), to namecheck other terms colonised by zombilingo. Team players play the game of excellence within what Cheek (2017) labelled the ratings rodeo. Further, as Saunders and Ramírez (2017) remind us, "since excellence is a measure of a thing, and since everything in post-secondary education is committed to excellence, everything must be measured" (p. 399). *Excellence R Us*.

Unfortunately, criteria for audit, performance and measurement are also colonised by corporatisation and its zombilingo about 'quality'. Lorenz (2012) noted that "the paradoxical and disastrous effect of the introduction of NPM, with its self-referential notions of accountability and quality, is that someone can be an excellent teacher and researcher and at the same time be assessed as poor by the QA system" (p. 619). Those in the neoliberalist, [New Public Management] NPM and zombilingo camps, clearly have a different construct of quality and its manifestation, excellence. Lorenz (pp. 618-619) writes:

Quality ... is concentrated upon systems and processes rather than outcomes. [Quality Assurance] QA is built on the assumption that any properly constituted organization should be based around a system of auditing systems and processes... The product of a QA system is therefore quality assured by definition—without necessarily guaranteeing its excellence or fitness for use.

The zombies are both those brainwashed, brain-eaten even, by zombilingo, and those forced to conform to its behaviours, often against their ethical well-being (Ryan, 2012), becoming anxious "nobodies" (Fleming, 2021, p. 116). Unsurprisingly, a zombie university (Smythe, 2017) produces zombie students (Ryan, 2012; Smythe, 2017). They are zombified because, with the failure of liberal arts and science ideology, they are made into work-ready agents of the market (Waller & Wrenn, 2023), which in zombilingo is often called 'social mobility' (Beighton, 2018). 'Work-readiness' as defined by the voices of employers in the echo Chamber of Commerce, usually means 'ready to be put to work in the money-making machine'. Automaton in service of the bottom line.

Considering that, 'excellence' is as non-neutral a term as 'quality' (Giroux, 2009), the neoliberalised "performative worker" is, as Ball (2003) wrote, "a promiscuous self, an enterprising self, with a passion for excellence" (p. 16). Whose excellence, we ask? That of the organisation, that of the performing educator or even that of their learners striving to get over the line or, perhaps, to find excellence. The performative, promiscuous worker is the brainwashed zombie, striving, with plenty of help from the neoliberal system, by the alpha in the "academic star complex" (Fleming, 2021, p. 116) in *Academic star wars* (Yudkevitch et al, 2023).

Ball (2003) is just one scholar who shows us that this promiscuity takes the form of wantonly reaping grants and outputs to release the university from obligations to fund research internally via salary; and to reward those who contribute most to such regimes as Australia's ERA (Australian government, Excellence in Research for Australia, 2023). Seen this way, research, under which higher degree research is subsumed, is an exercise of a *homo oeconomicus* identity than any authentic conception of 'excellence' or making a critical or transformative difference (Skea, 2021). As Roberts (2007) wrote of the New Zealand context: "Research is a competitive, self-interested, instrumental, outputs-oriented process" (p. 362). Roberts (2007) describes this species of 'performative' zombie with a reference to nationalism that reminds us of Swinburne's catch-call that we can make a difference together.

The ideal citizen... is a sophisticated, competitive, innovative and enthusiastic participant in the global economy, ever ready to apply what he or she knows (from research or other activities) to the goal of creating... a "prosperous and confident nation" (p. 363).

The neoliberalised subject is one that has bought into the rhetoric tying the work of the individual/*homo oeconomicus* to the patriotic 'national' (read *organisational*) good. But surely such individuals are docile, performing others' edicts? Elsewhere, Foucault (2008) defines the docile body as one that may be subjected, used, transformed and improved. While I am one who sees the assimilated *homo oeconomicus* as a zero, there are many who would praise the docile hero.

### The timeliness of excellence

One of the greatest impacts on postgraduate learners during the COVID period was on their timelines. The literature on postgraduate supervision has long emphasised the importance of timely completion as the most crucial outcome of the interplay of the three parties. Because time is money, a study of research supervision in Kenya used the word 'timely' 15 times (Noel et al., 2021); even in developing countries, it is priority number one. Yet educators know the virtue of timeliness emerges from the fortunate concatenation of a mentor's skill in co-negotiating, knowing, and managing processes, a learner's determination and discipline and the host organisation's ability to offer an environment that supports and ideally supplements both these things. It is a multi-stakeholder endeavour, but keeping to time is a responsibility allocated to the supervisor. Timeliness, though, primarily thrives with a preventative, interventionist and empathetic response from the supervisor or mentor (Manathunga, 2007) and not a punitive, top-down looming timeline. This ability, a major contributor to postgraduate educative excellence, is one of many in the toolbox of a good supervisor or a passionate mentor.

### The current study of excellence

This study is a work of professional practice and not traditional research. As such, it is grounded in real-world problems identified within the author's workplaces (Costley & Lester, 2012). The problem is how educators grapple with an ethos of 'just get them over the line' within a culture that professes a mission statement including a notion of excellence, albeit a two- (or even multi-) headed one. The gap it fills is not the result of an extensive literature search to find something relatively unexplored, but an observed and experienced disjuncture in real life informed by the Zeitgeist of literature critiquing neoliberalism. The gap I explore is that between the rhetorical, marketised concept of 'excellence' in higher education and the on-the-ground, gritty, real-world experiences of those closest to the phenomenon being explored.

This study is interpretivist: I re-present and, hence, interpret the experiential and human components. This interpretivist orientation evidences itself in my use of language, the recreation of mentor consciousnesses and the question of invested, privileged power perspectives in discourse about excellence. The method of evidence collection here is elicited narrative in response to the cue: *Share a narrative illustrating how you came to realise something important about what good mentoring looks like*. New mentors were asked to share stories of critical moments in their learning journeys. In the broader evidence set, I used thematic analysis, and as I did notice, the cliché *gets them over the line* used in three stories; this spurred my interest. The stories were recast to dilute/fictionalise identifiable references and to crystallise mentors' realisations about how excellence plays a pivotal role in their narratives. The participants agreed that I could use edited versions of their stories, substituted with critical incidents from my own story, bringing critically reflexive autoethnography to my entire paper (Ellis & Bochner, 2001). Safeguarding researchers' relationships with collegial mentors is an ethical issue applied seriously (Fulton & Costley, 2019). This study is ethics approved, Otago Polytechnic: HRE15-173.

In this paper and in my own professional practice, these stakeholders are postgraduate degree candidates and their mentors, and their experiences contain a shadow of the third responsible party: the host organisation. This paper applies tenets of narrative enquiry in that its 'truth' claim derives from the authentic stories of the lived experience of those close to the phenomenon over time, understanding how the individual and the culture are interconnected (Clandinin & Connelly, 1994). In my evocative autoethnography, I situate myself as a researcher within my study, often integrated with other ethnographic participants (Blanco, 2012). As in Hil's (2012) work, the social reality of the narrator presents a perspective on the object of enquiry, excellence, and this is what, in part, makes it critical. Clandinin (2013) claims narrative operates in the middle of an experience and should heed participants' temporality, sociality and place. The three narratives presented here are, thus, comparable and representative but hardly comprehensive.

The stories themselves comprise evidence (formerly known as data) and reflect and refract a speculative 'truth' set within a framework of my curation. Clearly, I write and curate from the closeness of my experience that Bochner and Ellis's (2016) legacy affords, and accordingly, I link "evocative personal narrative to cultural criticism" (p. 25) but narrow my scope to an examination of excellence. The phenomena I examine in this study are expressed in the title: the 'execution' of excellence, both in terms of how it is carried out procedurally, and in terms of how it is murdered by lexical slaughter. By way of a disclaimer, I need to say that no managers were harmed in the creation of this artefact, and their bottom lines remained untouched. The term 'excellence,' however, was. I now present the three professional masters/doctorate mentor narratives, each detailing a critical incident in their practice that led to a realisation of the nature of excellence.

## **Narratives of (not) excellence**

### **Narrative 1**

I inherited a 'legacy' Masters learner whose research work had involved implementing a [redacted] as a method of thinking about and planning a cultural event within the practicum of an undergraduate culinary arts programme. A lockdown in the weeks prior to the [redacted] stunted the potential of this approach, but the practical plans were underway and the event, which was the beating heart of the phenomenon under investigation, went ahead. Luckily, the date of this cultural event occurred after the end of the projected lockdown and for once lockdown was not postponed.

The learner had to rely on a retrospective, remembered narration. Now, looking back, we can see she could have/should have interviewed those learners involved in the process of planning to ensure a solid, time- and event-specific data set was preserved, but the existing ethics application did not cover such an approach and the committee had gone into COVID-inflicted recess. Hindsight is a fine thing, and when the work came to its final oral assessment, several other lockdowns later, the assessors were full of the might haves and could haves which, under normal circumstances, might have seemed reasonable. The candidate had done her best with severely limited retrospective case study data but had not been able to muster an autoethnographic approach since that also required forward-looking ethics. She grounded her work in the theory informing her original approach, but it was clear now that it did not fit epistemologically. The lack of fortune continued. Her supervisor was forced to resign. Her thesis was ready for review, but its methodology was mismatched to her proposal and her entire research direction had turned 360 degrees. Her reviewer told her that her work lacked rigour and methodological soundness, and this is the point at which I inherited her with a memo to please 'get her over the line'.

It was clear that her groundedness in her subject and practice was strong; that her passion had been burning, and that her initial planning had resulted in a viable and methodologically logical approach to addressing her enquiry. She strove for the 'excellence' that Hegarty (2011) relates to critical

reflection and cultural imagination in culinary arts education. But she had not prepared for the unknown. Although she had generated findings, discussed them and raised recommendations, the enquiry was, it was now evident, built upon sand. How did it get to this stage without anyone realising? I felt that a student whose natural inclination was towards excellence was being pushed just to pass, no more, because of forces beyond the control of herself, her mentors or indeed, the faculty. These forces were the demons of timeliness and its adjunct economic and KPI flow-ons. I felt that no matter how well I supported this candidate, she was caught between an advertised, promised excellence and the likelihood of achieving mediocrity. Her getting a mere pass was reflective of neither the organisation's promise nor the promise she had shown. As for myself as a stand-in, there was no possibility of achieving excellence; I was unable to make a silk purse out of a sow's ear, especially under time pressures. Reluctantly, she was eventually taken over the line, but it could have, should have, been so much better.

### **Narrative 2**

The second time my professional practice doctoral learner presented her thesis, our research management expressed the need for her to 'get over the line', that mythic liminal space in the borderlands where three examiners converge like confluent rivers. Hers had been a complex journey through rough landscapes of multiple COVID lockdowns and even the cessation of educational establishments in which her study, and certainly her evidence gathering, was grounded. She had engaged with literature and methodology and created an acceptable proposal with the potential for excellence. Along with COVID, life happened. Invited participants, initially keen, withdrew as their circumstances had changed, and due to the pandemic, they were no longer in the zone. Diverted by the pandemic, people were unable to dedicate time or energy to be interviewed even via Zoom. She pivoted and extended the sample participant group – twice – keeping checks with the ethics advisor as she did. Perhaps resilience and resourcefulness are true hallmarks of excellence in a time of unforeseen crisis? Perhaps creating a compelling narrative of the experienced messiness of research-in-practice was itself an artefact of doctoral rigour?

Data collection was suspended indefinitely, and other parts of the thesis were written. A professional as well as a candidate, she was made redundant, and the context of her enquiry disappeared. A year on, she assumed work elsewhere, so her work context changed. In the end, the data was eclectic but still told a strong, if messy, practice story, but, perhaps, was no longer totally congruent with what she had written. The final work was on time and authentic in terms of lived practice and evidence of critical reflection, but the choppy methodology and thinner-than-hoped-for findings narratives seemed mediocre to examiners. There was no space in the (externally quality assured) assessment criteria to assess for resilience and resourcefulness, and no possibility to reframe the work as a COVID-era best effort. There was a perceived line, and it wasn't over it yet.

Where the energies of learner and mentors had been strong, the thesis was the best it could be and, if treated as a portfolio, addressed the graduate outcomes competently. Not helped by MS Teams outages, the oral assessment seemed confrontational, and the candidate was intimidated into underperformance rather than supported to success. A new thesis was requested with additional data, but with all the time and energy expended, it was a tall order, and the data-gathering boat had sailed. What would have been and could have been potentially excellent was consigned to mediocrity. However, we need to reflect on where the mediocrity truly lies in this narrative and understand the network of factors behind the execution of excellence.

### Narrative 3

There was considerable difference between my perception, in my role as mentor, of my master's learners' success and that of the three assessors in the final oral examination. I had thought the work was sound, but not spectacular; thorough, but not as original a contribution to scholarship as it could have been. I wonder about the impact of multiple lockdowns and natural disasters on the data collection environment and on the workplace of this learner, which was, literally, washed away in a deluge following a major climate event. When learners are on the clock, and when they have clocked up every last possible extension, and life happens, and happens again, and happens repeatedly, I wonder how much that happened was within the learner's locus of control. Is there room for empathy as a component of excellence, or is keeping to time everything?

I wonder about the extent to which circumstances limited my learner's access to achieving a grade of excellence. I know this learner did have the appetite and enthusiasm for excellence, at least initially, but they lost heart because of the brickbats fate chucked, and then they received letters about the imminent ending date of enrolment. There was a hasty assemblage of materials, including data narratives from interviews long delayed, then relayed by Teams, whose transcriptions were a mangled mess. The final work was repeatedly edited but still exhibited signs of haste. The learner's final assessment viva was well-communicated, and the final assessment report was positive, but their final mark told them they were mediocre. The potential for excellence promised by the organisation, the programme and the mentors, who were, after all, close to the work, was countermanded by a score that said *no PhD enrolment prospect for you, sucker*.

### Discussion

These are three of many such narratives that demonstrate how the pandemic exposed the vacuity of higher education's promises of excellence within a neoliberalist ideology. The ideology has proven itself lacking in resilience and adaptation to change and void of the compassion that we might have thought a pandemic might have emphasised. The discourse of just getting learners over the line (or not, as in Narrative 2) is the utterance of those colonised by a neoliberalist culture forced to accept that mediocrity will

do. The facts that the stage was set for potential excellence, but that managerialism proved inflexible in the face of the unforeseen, indicate that the higher education organisations did not serve their learners with the excellence they might have expected. There were opportunities for the exercise of compassion, for refiguring assessment events impacted adversely by COVID in the name of natural justice and for seeing past the weaknesses occasioned by circumstances beyond the agency of either educator or learner. They are stories of excellence denied by the neoliberalist knee-jerk factors of timeliness and adherence to process.

Authentic excellence is a project of teamwork and includes mentorly passion and experience but its existence falters when every relationship in the team and beyond is, in Giroux's words, "ultimately judged in bottom-line, cost-effective terms" (2009, p. 673). All three stories may have played out more positively into spaces of authentic excellence without such ideological constraints as time always being money. There are cases where the exercise of compassion may sometimes lead us closer to excellence, thereby achieving more social and cultural capital in the long run. Looming over these narratives is a master narrative of COVID-resistant neoliberalism. Giroux (2009, p. 670) relates Leopold's (2007) conception of the entrepreneurial professor: educators must be trained "to watch the bottom line" and attend to principles of finance, management, marketing and brand identity in the common quest of a high-quality product, namely new knowledge. Watching that bottom line by merely getting learners over the line leads to mediocrity and nowhere near a high-quality product. This may be a management function, but it is not that of the educator and won't be embraced by the learner.

### Conclusion

This study has moved from the macrocosm of HE internationally, where the slippery, contested concept of 'excellence' has been hijacked and occupied by vested, particularly neoliberalist, interests, to three micro-narratives demonstrating how powerful exponents of *mediocracy* limited learner and educator access to authentic excellence in the wake of the COVID era. Apart from extending the finish line and improving assessment processes for those approaching it, there are other hopeful possibilities to mitigate against future mediocre zombiedom in the assessment of HE degrees. Regarding the university 'world' specifically, Hil (2012) suggested we must "routinely reframe language by referring to... community, public education, students rather than consumers, dialogue and debate rather than inputs outputs and impacts" (2012, p. 217). This may be achieved by "do[ing] away with intrusive monitoring and subsequent zombification of academics" and affording "a return to community, collegiality, fun, soul, and passion" (2012, p. 209). Hil advocates, in other words, authentic excellence, not neoliberal excellencespeak, and a culture of community over *homo oeconomicus*. That was in 2012, and in 2022, Hil (et al.) thought COVID might have catalysed change, at least in Australian public universities, but it has not yet.

Relanguaging to reauthenticate excellence away from *Excellence R Us* (Moore, et al., 2017) is one thing, but there is a need, too, to provide an impetus for creativity and criticality to abate what Bighton (2018) called *catatonia*, which echoes the vacuity Collini (2012) saw in neoliberalist discourses of excellence. As far back as 2008, Clegg (2008) had argued for the application of creative and critical life force, *Eros*, to HE as a counter to neoliberal catatonics. It is akin to what Tan (2022, p. 158) sees as “mindsight, attentive love, and storytelling”, which are components of authentic excellence. Through such humanist and posthumanistic approaches, emancipation from the hemmed-in non-human logic of neoliberalism might open spaces for agency, authentic resilience, shared stories and fresh possibility. Reinvesting in a bottom-up approach can be as simple as reconsidering what teaching well looks like and ensuring that professional development policies support strengthening cultures’ research: practice nexus and collaborative opportunities. A reflective, experience-led, democratised approach based on praxis enquiry and incorporating critical thinking characterises leadership in HE teaching (Brookfield et al., 2023). Re-establishing the educator as the leader of excellence and not an entrepreneurial professor enables teacher agency as an act of resisting neoliberalism. It goes some way to reframing educator agency and affording authentic excellence, viewing it as social and cultural capital and not just a fraction of the bottom line.

What HE failed to reframe was the cockroach hegemony of neoliberalism (Cerny, 2010) and the leadership model needed in such uncertain times (Balasubramanian & Fernandes, 2022). Mirowski’s fearful 2013 thesis about the undeath of neoliberalism in *Never let a good crisis go to waste* is transpiring, as Fleming et al. (2021) insinuate. It is embedded (Cerny, 2010). It is democracy’s nemesis (Giroux, 2009). COVID appeared to offer a promising sea change away from the neoliberal (for example, Connell, 2019; Healey & Barish, 2019; Andrew et al., 2020; Balasubramanian & Fernandes, 2022). Fleming (2021) articulated the hope colourfully: “beleaguered by managerial-bloat, business bullshit and a COVID-compromised economic environment, the idea of the modern university may soon come to an end” (p. 19) but admits the pandemic focussed “ugly truths” (Fleming et al., 2021, p. 111). However, as Garrick (2014) had written: “alarmingly, we so often appear to return to ‘business as usual’, as if nothing had really happened... rapacious corporate greed, avarice and corruption lurk ever close to the surface” (pp. 151-152). Thus, the bottom line is still at the top for HE’s colonisation by neoliberalism. So far, HE has missed the chance to implement the change-oriented reflections, hopeful thoughts, emancipatory actions and insightful research of many during this period (Fleming et al., 2021) and set in motion a change strategy step-by-step to leave neoliberalism behind.

## References

*The Age*. (2016, January 24). Bottom line always at the top for neoliberalism. <https://www.theage.com.au/national/victoria/capitalism-20160123-gmcme0.html>

Andrew, J., Baker, M., Guthrie, J., & Martin-Sardesai, A.

(2020). Australia’s COVID-19 public budgeting response: The straitjacket of neoliberalism. *Journal of Public Budgeting, Accounting and Financial Management*, 32(5), 759-770. <https://doi.org/10.1108/JPBAFM-07-2020-0096>

Andrew, M. (2022). Come to the cabaret: Voices from the modern university. *Journal of Applied Learning and Teaching*, 6(2), 17-27. <https://doi.org/10.37074/jalt.2023.6.2.19>

Australian Government (2023). *Excellence for Research in Australia (ERA)*. <https://www.arc.gov.au/evaluating-research/excellence-research-australia>

Balasubramanian S., & Fernandes C. (2022). Confirmation of a crisis leadership model and its effectiveness: Lessons from the COVID-19 pandemic. *Cogent Business & Management*, 9, 1-31. <https://doi.org/10.1080/23311975.2021.2022824>

Ball, S. (2003). The teacher’s soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215-228. <http://dx.doi.org/10.1080/0268093022000043065>

Barnett, R. (2013). *Imagining the university*. Routledge.

Bighton, C. (2018). A transversal university? Criticality, creativity and catatonia in the global pursuit of higher education excellence. In D. R. Cole & J. P N. Bradley (Eds.), *Principles of transversality in globalization and education* (pp. 47-66). Springer Nature. [https://doi.org/10.1007/978-981-13-0583-2\\_4](https://doi.org/10.1007/978-981-13-0583-2_4)

Blanco, M. (2012). ¿Autobiografía o autoetnografía? (Autobiography and Autoethnography). *Desacatos*, 38, 169-178.

Bochner, A., & Ellis, C. (2016). *Evocative autoethnography: Writing lives and telling stories*. Routledge. <http://dx.doi.org/10.4324/9781315545417>

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). Greenwood. <https://www.socialcapitalgateway.org/sites/socialcapitalgateway.org/files/data/paper/2016/10/18/rbasicsbourdieu1986-theformsofcapital.pdf>

Brookfield, S., Randolph, J., & Tan, S. (2023). *Teaching well*. Taylor & Frances.

Brusoni, M., Damian, R., Grifoll, J., Jackson, S., Kömürçügil, M. M., Matveeva, O., Motova, G., Piszcz, S., Pol, P., Roslund, A., Soboleva, E., Tavares, O., & Zobel, L. (2004). *The concept of excellence in higher education*. European Association for Quality Assurance in Higher Education. <http://dx.doi.org/10.13140/RG.2.1.2146.7683>

Canning, J. (2019). The UK Teaching Excellence Framework (TEF) as an illustration of Baudrillard’s hyperreality. *Discourse: Studies in the Cultural Politics of Education*, 40(3), 319-330. <https://doi.org/10.1080/01596306.2017.1315054>

Carpenter, C., & Ker, G. (2017). Learning from experience: Exploring professional identity growth through work-based

- learning processes. *Scope: Contemporary Research Topics (Flexible Learning)*, 2, 22–32.
- Cerny, P. G. (2010). Embedding neoliberalism: The evolution of a hegemonic paradigm. In P. G. Cerny (Ed.), *Rethinking world politics: A theory of transnational neopluralism* (pp. 128–156). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199733699.003.0007>
- Cheek, J. (2017). Qualitative inquiry and the research marketplace: Putting some +s (pluses) in our thinking, and why this matters. *Cultural Studies ↔ Critical Methodologies*, 17(3) 221–226. <https://doi.org/10.1177/1532708616669528>
- Clandinin, D. (2013). *Engaging in narrative inquiry*. Left Coast Press. <https://doi.org/10.4324/9781003240143>
- Clandinin, D., & Connelly, F. (1994). Personal experience methods. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 413–427). Sage.
- Clegg, S. (2007). The demotic turn – excellence by Fiat. In A. Skelton (Ed.), *International perspectives on teaching excellence in higher education: Improving knowledge and practice* (pp. 91–102). Routledge.
- Clegg, S. (2008). Creativity and critical thinking in the globalised university. *Innovations in Education and Teaching International*, 45(3), 219–226. <https://doi.org/10.1080/14703290802175982>
- Collini, S. (2012). *What are universities for?* Penguin.
- Connell, R. (2019). *The good university: What universities actually do and why it's time for radical change*. Monash University Publishing.
- Costley, C., & Lester, S. (2012). Work-based doctorates: Professional extension at the highest levels. *Studies in Higher Education*, 37, 257–269. <https://doi.org/10.1080/03075079.2010.503344>
- Cui, V., French, A., & O'Leary, M. (2021). A missed opportunity? How the UK's teaching excellence framework fails to capture the voice of university staff. *Studies in Higher Education*, 46(9), 1756–1770. <https://www.tandfonline.com/doi/full/10.1080/03075079.2019.1704721>
- Elias, J., & O'Leary, B. (2023). Where the public sees value in higher ed. *The Chronicle of Higher Education*. [https://www.chronicle.com/article/where-the-public-sees-value-in-higher-ed?utm\\_campaign=che-social-20231113&utm\\_medium=o-soc&utm\\_source=fb](https://www.chronicle.com/article/where-the-public-sees-value-in-higher-ed?utm_campaign=che-social-20231113&utm_medium=o-soc&utm_source=fb)
- Ellis, C., & Bochner, A. P. (2001). Autoethnography, personal narrative, reflexivity: Researcher as subject. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 733–768). Sage. [https://www.academia.edu/4551946/Autoethnography\\_Personal\\_Narrative\\_Reflexivity\\_Researcher\\_as\\_Subject](https://www.academia.edu/4551946/Autoethnography_Personal_Narrative_Reflexivity_Researcher_as_Subject)
- Fleming, P. (2021). *Dark academia: How universities die*. Pluto Books. <https://doi.org/10.2307/j.ctv1n9dkhv>
- Fleming, P., Randolph, J., & Tan, S. (2021). 'Never let a good crisis go to waste'. An interview with Professor Peter Fleming on dark academia, the pandemic and neoliberalism. *Journal of Applied Learning and Teaching*, 4(2), 110–120. <https://doi.org/10.37074/jalt.2021.4.2.14>
- Foucault, M. (2008). The birth of biopolitics (G. Burchell, Trans.). In M. Senellart (Ed.), *Lectures at the Collège de France 1978– 1979*. Palgrave Macmillan.
- Fulton, J., & Costley, C. (2019). Ethics. In C. Costley & J. Fulton (Eds.), *Methodologies for practice research: Approaches for professional doctorates* (pp. 77–92). Sage.
- Gillies, R. (2014). Cooperative learning: Developments in research. *International Journal of Educational Psychology*, 3(2), 125–140. <http://dx.doi.org/10.4471/ijep.2014.08>
- Giroux, H. (2009). Democracy's nemesis: The rise of the corporate university. *Cultural Studies. Critical Methodologies*, 9(5), 669–695. <https://doi.org/10.1177/1532708609341169>
- Giroux, H. (2014). Public intellectuals against the neoliberal university. In N. Denzin & M. Giardina (Eds.), *Qualitative inquiry outside the academy* (pp. 35–60). Left Coast. <https://doi.org/10.4324/9781315421339-1>
- Giroux, H. (2017). Neoliberalism's war against higher education and the role of public intellectuals. In M. Izak, M. Kostera, & M. Zawadzki (Eds.), *The future of university education* (pp. 185–206). Palgrave. [https://doi.org/10.1007/978-3-319-46894-5\\_10](https://doi.org/10.1007/978-3-319-46894-5_10)
- Giroux, H. (2019). *Neoliberalism's war on higher education*. Haymarket.
- Garrick, J. (2014). The limits of knowledge management in contemporary corporate conditions. *International Journal of Learning and Change*, 7(3/4), 141–155. <http://dx.doi.org/10.1504/IJLC.2014.063136>
- Goode, C. (2023). Rebels, rogues, and risk-takers: Insights into personal characteristics of national tertiary teaching excellence awardees. *International Journal for the Scholarship of Teaching and Learning*, 17(1). <https://digitalcommons.georgiasouthern.edu/ij-sotl/vol17/iss1/17>
- Grifoll, J. (2016). External quality assurance agencies and excellence in higher education. *Educación Médica*, 17(3), 94–96. <http://dx.doi.org/10.1016/j.edumed.2016.08.001>
- Gunn, A. (2018). Metrics and methodologies for measuring teaching quality in higher education: Developing the Teaching Excellence Framework (TEF). *Educational Review*, 70(2), 129–148. <https://doi.org/10.1080/00131911.2017.1410106>
- Healey, R., & Barish, J. (2019). *Beyond neoliberalism: A narrative approach*. <https://narrativeinitiative.org/resource/beyond-neoliberalism-a-narrative-approach/>
- Hegarty, J. (2011). Achieving excellence by means of critical reflection and cultural imagination in culinary arts

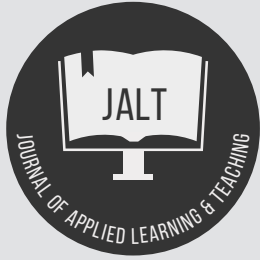
- and gastronomy education. *Journal of Culinary Science & Technology*, 9(2), 55–65. <http://dx.doi.org/10.1080/15428052.2011.580705>
- Hil, R. (2012). *Whackademia: An insider's account of the troubled university*. Sydney: NewSouth. <https://www.jstor.org/stable/j.ctt5vkpq8>
- Hil, R., Thompsett, F., Lyons, K., Joannes-Boyau, R., Lake, S., Lucas, A., McCallum, A., O'Connor, J., Pelizzon, A., Tregear, P., & Vodeb, O. (2022). Over the horizon: Is there an alternative to neoliberal university governance? *Social Alternatives*, 41, 1. <https://research-repository.griffith.edu.au/bitstream/handle/10072/426350/Hil6446151.pdf?sequence=2>
- Hilton, J. (1933). *Goodbye Mr Chips*. Hodder & Staunton.
- Katz, L. (2015) Against the corrosive language of corpspeak in the contemporary university. *Higher Education Research & Development*, 34(3), 554-567, <http://dx.doi.org/10.1080/07294360.2014.973377>
- Leopold, P. L. (2007). The professorial entrepreneur. *The Chronicle of Higher Education*. <https://www.chronicle.com/article/the-professorial-entrepreneur/>
- Lorenz, C. (2012). If you're so smart, why are you under surveillance? Universities, neoliberalism, and new public management. *Critical Enquiry*, 38(3), 599-629. <https://doi.org/10.1086/664553>
- Manathunga, C. (2007). Early warning signs in postgraduate research education: A different approach to ensuring timely completions. *Teaching in Higher Education*, 10(2), 219-233. <https://doi.org/10.1080/1356251042000337963>
- Mirowski, P. (2013). *Never let a serious crisis go to waste: How neoliberalism survived the financial meltdown*. Verso.
- Moore, S., Neylon, C., Eve, M. P., O'Donnell, D. P., & Pattinson, D. (2017). *"Excellence R Us": University research and the fetishisation of excellence*. Palgrave Communications. <http://www.palgrave-journals.com/articles/palcomms2016105>
- Noel, J., Wambua, B. K., & Ssentamu, P. N. (2021). Invest in research supervision, enhance timely completion of postgraduate studies. *RMC Journal of Social Sciences and Humanities*, 2(1), 35-47.
- Owens, A., Daddow, A., Clarkson, G., & Nulty, D. (2021). What is the price of excellence in learning and teaching? Exploring the costs and benefits for diverse academic staff studying online for a GCHE supporting the SoTL. *Teaching & Learning Inquiry*, 9(1). <http://dx.doi.org/10.20343/teachlearninqu.9.1.12>
- Roberts, P. (2007). Neoliberalism, performativity and research. *International Review of Education*, 53(4), 349–365. <http://www.jstor.org/stable/27715397>
- Ryan, S. (2012). Academic zombies: A failure of resistance or a means of survival? *Australian Universities Review*, 54(2), 3–11. <https://eric.ed.gov/?id=EJ981188>
- Saunders, D. B., & Ramírez, G. B. (2017). Against 'teaching excellence': Ideology, commodification, and enabling the neoliberalization of postsecondary education. *Teaching in Higher Education*, 22(4), 396–407. <https://doi.org/10.1080/13562517.2017.1301913>
- Skea, C. (2021). Emerging neoliberal academic identities: Looking beyond homo economicus. *Studies in Philosophy and Education*, 40(4), 399–414. <https://doi.org/10.1007/s11217-021-09768-7>
- Skelton, A. (2004). Understanding 'teaching excellence' in higher education: A critical evaluation of the national teaching fellowship scheme. *Studies in Higher Education*, 29(4), 451–468. <http://dx.doi.org/10.1080/0307507042000236362>
- Skelton, A. (2005). *Understanding teaching excellence in higher education: Towards a critical approach*. Routledge. <http://dx.doi.org/10.4324/9780203412947>
- Smythe, J. (2017). *The toxic university: Zombie Leadership, academic rock stars and neoliberal ideology*. Palgrave.
- Tan, E. (2022). 'Heartware' for the compassionate teacher: Humanizing the academy through mindsight, attentive love, and storytelling. *Journal of Applied Learning and Teaching*, 5(2), 152-159. <https://doi.org/10.37074/jalt.2022.5.2.ss1>
- Timar, T., & Kirp, D. (1988). *Managing educational excellence*. Falmer Press.
- Tinto, V. (2017). Through the eyes of students. *Journal of College Student Retention: Research, Theory & Practice*, 19(3), 254-269. <https://doi.org/10.1177/1521025115621917>
- Tracey, S. (2010). Qualitative quality: Eight 'big tent' criteria for excellent qualitative research. *Qualitative Inquiry*, 16, 837-851. <https://doi.org/10.1177/1077800410383121>
- van Staveren, I. (2023). The paradox of resilience and efficiency. *Journal of Economic Issues*, 57(3), 808-813, <https://doi.org/10.1080/00213624.2023.2237861>
- Waller, W., & Wrenn, M. V. (2023). The COVID-19 crisis as an opportunity to (further) extend neoliberalism into the higher learning. *Journal of Economic Issues*, 57(3), 814-828, <https://doi.org/10.1080/00213624.2023.2237862>
- Watson, D. (2003). *Death sentence: The decay of public language*. Vintage.
- Wood, W., & Su, F. (2017). What makes an excellent lecturer? Academics' perspectives on the discourse of 'teaching excellence' in higher education. *Teaching in Higher Education*, 22(4), 451-466, <https://doi.org/10.1080/13562517.2017.1301911>

Woudstra, A., & Powell, R. (1989). Value chain analysis: A framework for management of distance education. *The American Journal of Distance Education*, 3(3), 7-21. <https://doi.org/10.1080/08923648909526675>

Yudkevich, M., Altbach, P. G., & Salmi, J. (2023). Academic star wars: Excellence initiatives in a global perspective. *Massachusetts Institute of Technology*. <https://doi.org/10.7551/mitpress/14601.001.0001>

Copyright: © 2024. Martin Benedict Andrew. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## The Great Resignation: The simple joys of not belonging

Martin Benedict  
Andrew<sup>A</sup>

A

*Associate Professor of Work-Based Learning, Otago Polytechnic, Dunedin, Aotearoa/New Zealand*

---

### Keywords

Academic identity;  
belonging;  
neoliberalism;  
not belonging;  
public intellectuals.

---

### Abstract

As the 2020s march on into a post-COVID age, an increasing trend for academics to exit their current academic positions or to leave academia altogether can be observed internationally and locally. Consequently, a sizeable body of experts accessible to higher education but geographically beyond its ivory towers and psychologically outside its neoliberal grip has come to exist. These para-academics and public intellectuals continue to contribute to communities of teaching, learning, and researching but do so often without affiliation. This study explores the relational link between the archaic notion of affiliation and what it means to 'belong' to a university as staff. The study problematises belonging as an assimilative designator of an organisation's culture and suggests that belonging, as employed in teaching and learning discourse, as a trust-based mode of building community, is a different beast than that conceived by neoliberal universities. Using vignettes as narrative enquiry, the paper retells and curates six accounts of academics making transitions out of academic positions and finding fresh educational contexts for belonging. These emancipatory narratives move through spaces of trauma into authentic places of reclaimed identity, most notably as independent public intellectuals within a broader context of global citizenship. The narratives show us what life after being academically affiliated can look like when individuals exercise critical resilience to establish academic identities beyond the neoliberal university.

---

### Correspondence

[martin.andrew@op.ac.nz](mailto:martin.andrew@op.ac.nz)<sup>A</sup>

---

### Article Info

Received 6 March 2024  
Received in revised form 21 March 2024  
Accepted 22 March 2024  
Available online 15 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.36>

## Introduction

Varying degrees of not belonging to academe, universities, institutions, or faculties appear in accounts of tertiary educators' departures within the research sub-genre known as "the toxic university" (Smyth, 2017) or "dark academia" (Fleming, 2021) and in Barcan's (2013) work on why academics leave. My study explores the neoliberal underpinnings of 'belonging' and troubling it to suggest the category of 'not belonging' ('Un-belonging' is used in studies with learners). Not belonging also challenges the notion of 'affiliation' to describe an educator's identity in academic orbits, such as those of conferences or professional organisations, as outdated and hegemonic. I propose that Giroux's (2014) notions of "public intellectual" and the concept of "para-academic" (Withers & Wardrop, 2014) are appropriate signifiers in this space. These titles align with "relational being" (Graham & Moir, 2022) as a more agentive description of authentic academics in a "supercomplex" world (Barnett, 2012) or, applying the updated COVID-era version, a VUCA (volatile, uncertain, complex, and ambiguous) world of overlapping wicked problems (Stein, 2021).

We may still 'belong' as global citizens with integrity despite geo-physically operating beyond the contracted walls of the hallowed university. If you 'belong' to an organisation, you are seen as 'affiliated' to it; yet a sinister sense of 'belonging', that of contracted ownership, a Faustian bargain, lingers within the connotations of 'belonging'—the moment when Mephistopholes 'belongs' to Faust. In this sense, 'belonging' implies an individual accords with the assimilative designators of an organisation's culture: its mission and core values, its employee behaviours, and world view. Invisible in 'affiliation' is any sense of an organisation's ethic of care, as may be the case with casualised workforces and precariats. This study suggests ethics (or 'duty') of care is increasingly replaced by "academic incivility" with its "bully culture" (Twale & DeLuna, 2008). How many academics happily sign work contracts in the knowledge that they sign themselves into a Faustian bargain with a neoliberal twist?

Narrative and autoethnographic studies of the lived experience of academic identity in anxious times, spearheaded by Sparkes (2007) and Poulos (2017), offer empirical yet visceral accounts of damaged subjectivities of individuals whose senses of belonging have gone (Andrew, 2020; Fleming, 2021). More often, they are shown as having been snatched away by the machinations of technocratic corporatism characterised by surveillance culture (Ball, 2003, 2012; Shore, 2010). Particularly significant is the "overloading of responsibilities" (Shore, 2010, p. 20) as workload intensification is disguised as dutiful service or lifelong learning. Indeed, Taylor (2013) pointed out that the use of business-oriented euphemisms such as "flexible delivery", "lifelong learning", and generic "excellence" obscure a harsher reality. Exercises of university brand-building and rebranding replicate this discourse, which Barnett (2012) regarded as forms of imaginative, ideological, and ethical constraint.

Drawing on the method of presenting multivocal narratives of lived experience, themselves based on elicited writing, this study suggests that not belonging for academics

in the sense of being independent/unaffiliated offers a constructive and rewarding possibility for higher education workers who, for such reasons as redundancy, resignation or the expiration of any honorary status, are no longer affiliated to a single master or are free from an institution that affiliates them. The enquiry addresses the question: *In what ways does not belonging to a tertiary institution enable and support independent academics?*

## The Great Resignation

As I was collecting data in the form of narratives of exiting tertiary education, as part of a broader project on changing academic identity in Australia and New Zealand, I discovered a rich and authentic data source in the US-based Facebook page "The professor is out", where departing/departed academics share their stories: those thinking of leaving solicit advice, doctoral learners, who realising that academia has no future, seek corroboration for their hunches, and academics in new roles express their regret that they stayed in academia as long as they did. Owan et al. (2024) wrote of the need to balance out the instrumental neoliberalist concept of "metrics" by fostering a "culture of rigorous and unbiased evaluation in the academic community" (p. 9). Interestingly, however, they methodologically base their study on evidence from a Facebook page, "Reviewer 2 must be stopped". This is a place where authors share mostly shocking experiences of the review process, demonising the ever-mean Reviewer 2. What I do present are six donated narratives of reformed academics, and their reflective stories echo core themes of the literature of "The Great Resignation" (Flaherty, 2022): being worked to death, bullying managers, losing positions in restructures, and realising there is more to life than *this*.

Barcan (2013) and Flaherty (2000) have identified a trend among scholars and educators, especially those later in their careers or post-PhD, who are part of the literature on "The Great Resignation," which refers to the significant departure of professors in the 2020s. They feel a sense of 'not belonging' within academia yet take on the role of public intellectuals, adopting an activist stance that challenges traditional academic identities based on institutional affiliations. This position underscores that aspects such as collegiality and the pedagogical process itself are integral to social justice (Goodall, 2010). It is also an adaptive position in that it embraces multiplicity and resists the capitalistic ideologies that infuse the hierarchal designation of status within organisations. The notion of academic 'affiliation' is a dinosaur of a feudal age that has passed; it meant adoption as (or possibly suckled as) a son in its medieval Latin cognate, figuratively extending to adoption by a society, a relationship, a belonging. The name 'academia' itself suggests a scholarly Platonic Grecian idyll, like 'Arcadia', from a golden but remembered age of nostalgic fiction (Tight, 2010). Any analogy between a university and a family or a community belongs, too, to a bygone age.

The critical and radical types of global citizenship education (GCE) discourse (Stein, 2021) remind me of the utopian days of collegiality (Tight, 2010), which Hil et al. (2022) hope can be reclaimed as part of a global rejection of

neoliberalism. It reminds me, too, of the Greek origins of 'cosmopolitanism' as *not being affiliated* with any given city-state or polis (Koukouraki, 2020). The idea of being a citizen of the world breaks down the exclusionary boundaries, perhaps best epitomised by the Trumpian wall, in favour of a cosmopolitan view of global citizenship more akin to another transformational utopian idea, the melting pot, the Confucian great unity (大同 / dàtóng), where we may all be family (Koukouraki, 2020).

Aotearoa/New Zealand, moving to a co-governance structure in its higher education organisations in the mid-2020s, is the only country to use the term *whanāu* ('family') with any degree of success, but even then, it can feel like imposed belonging. This is arguable because it ontologically clashes with the concept of *whakapapa* or heritage that academics in Aotearoa/ New Zealand figure into rote-learned *pepeha*, or individual, personal origin stories. *Pepeha* figure at the opening of meetings as part of the process of *mihi* (greeting) and can be figured as a *tangata whenua* (people of the land) or *manuhiri* (guest, visitor). Our varying stories of whakapapa show that we cannot truly be a blood *whanāu* but in a post-*Tiriti* (Treaty of Waitangi) sense, we can be a group with a common sense of endeavour, a community of practice. We all have a sense of belonging to our *tūrangawaewae* (our place to stand). While Aotearoa/ New Zealand seemed in 2023 to be moving towards a family-focussed sense of co-governance with its promise of unity as opposed to division, the populist hate rhetoric of the COVID era intersected with mark2-neoliberalism in an elected coalition government late in 2023 and 2024. For academics still passionate about learning and research but having no affiliated place to stand, the identities of the para-academic and public intellectuals offer much potential, and, even more, hope.

## Literature review

I have chosen to cover key sub-topics in this review. The first of these delves into the literature of belonging in anxious times (Press et al., 2022) under the heading 'Troubling belonging'. In the next section, I extend this sense of troubling to communities of practice as somewhat utopian sites where people ideally behave well. Finally, I consider what it is that academics leaving their formal positions are, in fact, leaving. They may, in fact, be leaving a vision of the university that had long since ceased to exist.

## Troubling belonging

Not all scholars support 'belonging' and its assimilative ideology as an ideal for students and, indeed, academics (Press et al., 2022, in a special issue entitled 'Pedagogies of belonging in an anxious world'). Few studies consider 'belonging' from a worker's perspective, the majority examining learner belonging as in the abovementioned special issue; or they afford the fostering of belonging, especially via designed collaborativism and evidence of instructor presence in online communities of practice, in both pre-COVID and COVID era settings (e.g. Andrew, 2024; Stafford, 2022). In a rare exception, Mulrooney and Kelly

(2020) demonstrated that 'belonging' was a value of import to students and staff alike and related to motivation and attainment. However, for staff, belonging was seen as "the degree of alignment between the role within the organisation and the personal needs of the employee" (Mulrooney & Kelly, 2020, p. 23). With a definition of 'belongingness' as "the congruence between the expectations of the role within the organization and the personal needs of the employee", Brion's (2015, p. 5) study of teacher morale is rare in its linking teacher morale, an asset of well-being, to belongingness, rationality, and identification.

Universities are not places of inclusion and belonging for all learners, particularly priority, non-traditional, mature-age, first-in-family, and low socio-economic background learners (Antonsich, 2010; Lähdesmäki et al., 2016; Crawford et al., 2022). Indeed, Kahu and Nelson (2018) stress that "viewing belonging as the outcome of both institutional and student factors recognises that belonging can manifest differently for each student depending on their background, their personality and other aspects of their experience." (pp. 65-66). Berryman and Eley (2019) call upon a more responsive and relational pedagogy to counter the "racial microaggressions and lateral violence" (p. 19) experienced by priority students, often indigenous, undermining their possibility of belonging. We trouble 'belonging', too, by questioning the neoliberalist hegemonic discourse of the individual, with individual academics scrambling competitively for metric points (Owan et al., 2024). The focus on individual performativity normalises for academics "isolationist, self-interested individualism at the expense of more collectivist, community-oriented ways of being in the world" (Press et al., 2022, p. 4).

The dominant pro-belonging view is represented by Healey and Stroman (2021), whose detailed work demonstrates that "building learning environments that support belonging, and therefore learning and well-being, for every student entails both challenging exclusion and promoting inclusion" (p. 9). They are spot-on, but this needs to occur within a relational, collaborative approach involving "care for and valuing of students as complex, situated, knowledgeable beings in their own right" (Graham & Moir, 2022, p. 12). This is because, according to Yuval-Davies's explanation,

belonging can be an act of self-identification or identification by others, in a stable, contested or transient way. Even in its most stable "primordial" forms, however, belonging is always a dynamic process, not a reified fixity, which is only a naturalized construction of a particular hegemonic form of power relations (2006, p. 199).

While fostering a sense of belonging is positive, particularly for a world reeling from a COVID-19 online meltdown, it feels different when management, "a particular hegemonic form of power relations", does it. Graham and Moir (2022) see institutional fostering of belonging among diverse learners as leading to "a culture of conformity and assimilation which perpetuates the injustices of those unable, or unwilling, to 'belong' due to their personal backgrounds, beliefs, or material circumstances" (p. 2). They argue that the notion of 'belonging', as a function of aspiration and allegiance,

is a *sous*-marketing exercise in “institutional retention and economic advantage” (p. 22), suggesting an enforced, non-critical conformity to an institution’s ideals, behaviours, and missions.

Similarly, when academics begin employment at any university, there are orientations involving inculcation in organisational policies, missions, and multiple in-person gatherings to meet other novices in the ‘cohort’. There are endless mandated tick-box online training about facets of the organisation’s culture including health and safety, cultural safety, bullying and harassment, rainbow rights, and whistleblowing. Coming to belong is a process of assimilative acculturation, but for both the learners and staff, as Graham and Moir (2022) discuss “the ways in which it can be enacted, and the motivations for doing so, stem from prevailing neoliberal agendas which seek to instrumentalise education and, in doing so, favour the experiences and values of dominant groups” (p. 12). *You will* belong, or else. You will belong – to the organisation that pays you.

### **Troubling belonging to communities of practice**

When people form a community to which they can belong, they “come together because they are able to identify with something—a need, a common shared goal and identity” (Hung & Der-Thang, 2001, p. 3). In this vision of ‘community’, be it a group, a faculty, an institution/ university, or the dream of academia, shared needs, goals, and identity are the key factors fostering belonging. These factors align with Wenger’s (1998) community of practice theory. Here, the imperatives of mutual engagement (the regular interactions of community members), joint enterprise (members’ common endeavours, goals, visions), and shared repertoire (ways of thinking, speaking, expressing, and remembering common to the community) unite eclectic individuals into a whole, where novices are supported by the experienced. Gaining any sense of community begins with feelings of membership, an affective, engaged, invested sense of wanting to belong (Wenger, 1998).

What happens, though, when the need is no longer mutual, the goal no longer common, and the shared identity too remote or ideologically alien to enable the maintenance of individual ethical authenticity? More specifically, what happens to university academics (and their learners) when the university endeavours which they have invested in no longer accord with that of the technocratic agendas of the new order? In this order, Ginsberg (2011) reactivates a zombie apocalypse with armies of functionaries, name-checked as “vice presidents, associate vice presidents, assistant vice presidents, provosts, associate provosts, vice provosts, assistant provosts, deans, deanlets, deanlings, each commanding staffers and assistants – who, more and more, direct the operations of every school” (p. 433). This parade depicts identities trapped in “a pre-determined paradigm of capitalist domination” (Neary & Wynn, 2016, p. 410). The technocracy-heavy structure of the modern university is at odds with the primary endeavours of teaching, learning, research, and social good.

What happens to us when we no longer share the same sense of community? We might go underground with like-minded colleagues to maintain what’s left of our integrity and write a book or a journal special edition, exemplified by *Social Alternatives* (2022) — *It’s time: the reform of Australian public universities as a strategy of collective solidarity* (Hil et al., 2022). We might turn to post-structural social critical theory which teaches us in an age of *vive la différence* to think about the identity of education workers as being beyond academic identity, offering increased possibilities for malleability and multiplicity in the spaces of the public intellectual (Neary & Wynn, 2016). To paraphrase Neary and Hagyard (2010), it is necessary to imagine and realise new forms of social institution for higher education founded on a fresh understanding of social capital as an abundance of knowledge, rather than the idea of education as a commodity (Neary & Wynn, 2016). This economy of social wealth offers fresh possibilities for academics as producers who don’t ‘belong’ and are not even emeritus or honorary. They may use their abundant knowledge for a public, not primarily institutional, good. Another strategy to regain agency and identity is to adjust our sense of belonging. Community of practice saw the sense of community as a psychological concept, comprising one’s sense of place, its people, their collegialities, their shared compassion, and their sense of belonging. When the sense of community erodes, so too does the sense of belonging.

### **Not belonging to what?**

Any number of scholarly names can agree with the assertion that lack of agency, powerlessness and anxiety increasingly denigrate academic identities up to and into the 2020s. This is an age when ninety percent of UK academics are cited as being ‘very unsatisfied’ with management (Fleming, 2021). The list may open with Slaughter and Leslie’s (1999) germinal critique of globalised, marketised academic capitalism and Tight’s early anthology of narratives (2000). The former foretells an environment of contradictions where faculty and professional staff expend their human capital stocks in increasingly competitive situations. The latter is the first of studies detailing how socio-political change impacts academics’ lived experiences, a theme presented in an age of increased anxiety by Loads et al. (2016), Poulos (2017), and Evans and Nixon (2015), who speak of the long shadow that European neoliberalism casts on colonies such as Australia.

Ball (2003, 2012, 2015) and Shore (2008, 2010) are leaders in presenting perceptive, emotive ‘Zeitgeist’ accounts of anxious life in the neoliberal university and they are compelling, authentic, and elegiac. Academe is a place where your soul becomes the property of the affiliated university (Ball, 2003) and where traditional community relationships of trust and professionalism have irreconcilably eroded (Shore, 2010). In a long catalogue of reasons why academics put in the hours, overwork is ultimately a badge of courage (Acton & Glasgow, 2015) in the performance of duty. Yet Ball (2003) declared: “Performance has no room for caring” (p. 224). Academics are subject to horrid Orwellian technologies of evaluation and audit, rendering academe itself as damned as Faust sucked dry by Mephistopheles (Ball,

2012). Gill (2009), Hil et al. (2022), and Tregear et al. (2022) curated narratives covering exhaustion, stress, overload, insomnia, anxiety, shame, aggression, hurt, guilt, feelings of out-of-placeness, fraudulence, and fear of exposure within the contemporary academy. Scholars of imposter syndrome will also recognise these feelings (Dews & Law, 1995). Berg et al. (2016) named the drivers of economic efficiency and intensifying competition as core sources of anxiety in Northern European academia, and life generally; these anxieties today manifest in the cost-cutting culture of degrowth. These values of outsiders from the marketplace have no place in academe. Withers and Wardrop (2014) observed, “scoundrels have infiltrated the academy—bureaucrats, managers and marketing ‘experts’—some of whom know very little, or even care about, education” (p. 6).

Giroux (2017) sounded a call to war: academe versus neoliberalism, but there is hope in his notion of ‘the public intellectual’ (Giroux, 2014): the independent thinker interrogating the text from the margins, or the independent thinker with critical resilience (Bottrell & Keating, 2019). Loads et al. (2016) share many stories from those struggling with performativity and intensification to those reconfiguring productive terrains. These (and other) studies are powerful, essential snapshots, and their visceral language is evocative and innovative, figuring the identity of academic discourse itself as evolving. Strategies for resisting neoliberalist ideology are relayed (Bottrell & Manathunga, 2019), but these are still largely from within academia and focus on maintaining self-integrity and learner-centredness despite authors being mostly still part of the machine and still complicit. They may be Barker’s (2017) zombies and nervous wrecks, hanging on obliviously and/or anxiously, or they may, like Barker, figure ways to negotiate the new terrain.

## Methodology

The vignette, also known as a scenario or situation, is a short story with characters who may be hypothetical or fictionalised and is used to prompt and elicit participants’ perspectives on difficult topics. In social science research, vignettes are used as a mode of data elicitation (Kandemir & Budd, 2018); but my use of ‘vignette’ describes how the narratives are collected and presented. Vignettes are an allegorical method of gaining narrative data from participants or an ethnographic mode of retelling stories (Bottrell & Keating, 2019). In this use of vignettes, the ethical imperative is to protect identities. As a narrative method of sometimes visceral representation, vignettes are autoethnographic, and may as such, draw on respondents’ partial happenings, fragmented memories, echoes of conversations, and corridor whispers (Sparkes, 2007). Poulos (2017) calls autoethnography a methodology of “resilience, resistance, and remembrance” (p. 1). He reminds qualitative researchers under fire in the academy and beyond that “we meet resistance with resistance, reproach with resilience, and disregard with remembrance” (p. 1). Such a relational narrative approach views an act of coming to know as a human and cultural construction (Polkinghorne, 1997). The narrative vignettes presented here are stretches of memory that testify to critical resilience and speak to the theme of

not belonging.

These narratives are stories from a wider, ethics approved project on the lived experience of members of academic communities past and present during turbulent times (Ethics: Victoria University, Melbourne, HRE16-204). I want to be clear that no cited evidence comes from ‘The professor is out’, ‘Reviewer 2 must be stopped’ or any additional source. A series of narratives problematising ‘voluntary’ redundancy has already appeared (Andrew, 2020). Taking a similar narrative approach, this paper problematises ‘belonging’ as it pertains to academics and their affiliation or psychological alignment with universities to which they may or may not have belonged, peripherally or centrally. All the narratives used in this particular study are from people in their 50s or very early 60s, from three men and three women. The narratives all come from the Australia-New Zealand region, with Narrative 1 extending into an Asian nation and Narrative 3 referencing time in the United Kingdom.

Participants were purposively sampled from the networks of the author and his co-researcher, and the resonance of the subject led to a snowball effect. Participants were asked to write a short reflective piece on one or more of a sequence of bullet-pointed themes related to loss of academic identity, one of which was belonging to the modernised university. These narratives have been curated to remove shadows of recognition and returned to the writers for confirmation. Each of them presents a vignette of transition, an academic identity in progress, and they all find a way to portray authentic critical resilience in the face of managerialist narratives of nihilistic resilience. This narrative process involves in each case a grappling with belonging, or not.

### **Narrative 1 (New Zealand and Asia): Change (mis) management—where can we belong?**

Leaving my work after 24 years was scary and liberating at the same time. I had to deal with two contradictory emotions; one was a miserable, dark sinking feeling, and another an elevating feeling that I will be free from this misery. To find myself without work was devastating. I lost my income, my ‘social status’, which was ‘guaranteed’ by being ‘institutionalised’. And I loved my job. I loved co-creating new knowledge with my students, my research, working with like-minded people, the energy of the place, and the opportunity to connect with local communities. There was so much that I loved about my work, which became a part of my identity. How could I be ‘me’ without it? However, the place had changed so much and I realised it was no longer the place that now existed only in my imagination.

At the same time, I was caught by this strong feeling of liberation that was coming from the realisation that I did not need to return to the place that made me sick over the last two or three years. The endless and mostly meaningless restructuring resulted in bringing the institution to its knees with the help of a new group of managers who’d appeared from nowhere. They couldn’t manage and they couldn’t lead due to an absence of future vision and a general lack of organisational history, combined with a lack of understanding of the academic environment and the

needs of the sector. Preoccupied with their own survival and keeping their jobs as restructuring continued, the new managers mastered *discipline and punishment*. Control over the dissemination of knowledge was only one of their skills. One day my colleague said to me bending over the coffee machine in the kitchen: 'We don't really know what is happening in our department, do we? Unless we were in the elite circle'. These words stayed with me because a week later, he passed away from heart failure. The selected circle comprised the fast-growing group of middle managers that included contractors; some had only recently graduated. It was not even the classic divide-and-conquer rule but a survival tactic of an inexperienced management who knew that they needed to surround themselves with sycophants who would feel their obligation to them and always support them. Darwinism at its worst.

I started to observe bullying around me. It was surreal. I knew that bullying existed, but I was lucky or naïve, perhaps, to have endured it without experiencing or observing it myself. Colleagues, respected academics, and experienced teachers tried to speak their minds, to raise their voices, critically but not with criticism. Everybody saw the need for change and 'wanted to be part of the solution,' using management jargon. Colleagues asked caring questions about the new direction of our programmes, our industry, our department, and our institution. The new managers felt threatened by the questions; they had no answers. Instead of initiating constructive dialogues, they initiated disciplinary measures. The new 'middle managers' simply shut down questions and reported misbehaviour. By being obediently silent, you have a chance to survive.

There was nothing left for many of us. People started to leave. Many left without securing a new job. My colleagues were leaving because they realised that staying much longer would affect their physical and mental health. Others were afraid to leave or could not leave because of their financial circumstances and many other reasons. Then one day, I faced the feeling that I didn't want to go to my office. I felt palpitations. I had this debate going in my head – *I'd loved my job, but I couldn't stay any longer. However, I had no plan B. My family is here. Should I try to hang in, perhaps? How long for?* This cacophony of voices in my head left me divided. An a-ha moment: continuing in this environment not only affects me physically, but also mentally. Time to go!

There is a saying: *when one door closes, another opens*. Things started to happen quickly with redundancies on offer. I signed a contract with another university a month before my final day. It was a part-time, temporary contract but a breath of fresh air, and psychologically, it was important to prove that I could do it. On arriving, I felt belonging—a good sign! I seriously thought of moving away from academia at the time but another job in the international university found me or I found it. I think I am in control of my life again and this is an empowering feeling. The last few years in my previous workplace became a distant memory, which I hate revisiting. I wish I could obliterate it completely; but as long as I can park it on memory lane, I am content.

## **Narrative 2 (Australia): Finding spaces for the public good outside the academy**

For some strange reason, I retain the view within the murky morass of neoliberal capitalism, that an academic is a person of high probity who works at a university and is committed to the pursuit of knowledge through an integrated approach to research, teaching, and professional and community service. For new knowledge to be forged in either tentative or more substantial forms, the work of academics must be carried out with autonomy and integrity, not at the whim of others. Of course, academics do not work in a vacuum but construct their activities around the principles, protocols, and codes of conduct that should have been established by the profession over long periods of time. Further, it is important that tertiary institutions have competent management and administrators to ensure that all aspects of university life can continue with efficiency and appropriateness, provided that management and administration do not interfere with the conduct of academic assignments. Unfortunately, the dominance and distortions of market forces over recent years, has determined that this separation of powers is often a distant memory. Educational quality, indeed, the honour and nobility of higher education, is at stake.

It may be possible to seek a breathing space, perhaps even to strengthen academic work, by establishing an independent existence, outside of the university environment. That is free at last, free at last to concentrate on working with a small group of research students, to write for a range of publications, to engage in various projects when available, and to undertake other educational and research activities for personal interest and satisfaction. Some formal contact with a host university will most likely be required. However, most of the incessant meetings and administrative tasks that face academics every day, would be eliminated. In other words, one might become a true, autonomous, professional academic. For the public good.

In many respects, these are moral decisions, of determining where the most good can be achieved, for the majority of participants. Establishing an independent office reduces contact with larger numbers of students and lessens the opportunity for all those informal discussions with staff on ideas and projects that excite. Wandering down the corridor for a chat is often when the ideas ferment. Financial considerations may mean the continuing necessity to write grant applications similar to the pressures of formal academic employment. Neoliberal dominance has made professional academic life in its truest guise very difficult to achieve and therefore can generate enormous frustration for those who want an honest academic relationship with knowledge, students, and colleagues. If it can be arranged, an independent, academic existence has many attractions.

## **Narrative 3 (Australia): Why flying under the radar is a useful strategy**

The three universities I have worked at are all in the same city. They range from a research-intensive, high-status university, a middling one and another that is of lower status. I completed a PhD at the first and worked there for eight

years in total. The bullying became intolerable and with the support of the union, I won a formal case, but nevertheless had to leave. When I was offered a tenured position at the middling university, I was very pleased and threw myself into teaching and research. A new head of school was appointed but turned out to be a sociopath. I was groomed and toyed with over a dark period of two years. I met every impossible target, and the administration load was so heavy that I rarely stopped working. I could not allow myself to buckle. It was an identity thing. Again, the union assisted when I formalised my accusation of bullying, which escalated to an accusation of victimisation. That means the person claiming bullying receives even more abuse for speaking up. The person from the Human Relations Department backed the manager and appointed a tame external consultant to examine the case. I lost. This level of bullying leads to a breaking point.

After leaving that tenured position, I was offered a five-year research-only role in a new research institute. I couldn't believe my good fortune. My earlier experiences influenced my decision to consciously fly under the radar, instead of joining in with the new university community. Research-focused positions are quite rare and highly competitive, and I felt animosity from colleagues in the faculty, who had high teaching and administrative loads. Given my outgoing nature, it was a considerable restraint not to join in. I resolved to work on this until it became my professional persona. Over time, I became more and more solitary as the focus became writing grant applications and publications and supervising PhD students – rather than working collegially with others. Over the eight-year period I spent at that university, I became friendly with half a dozen people, but few of them have been to my home. Reflecting on this significant change in professional identity has revealed just how strange it has been to continue to exist professionally, as an absent presence. Sanctioned bullying is widespread in universities, and it messes with people's lives.

#### **Narrative 4 (The United Kingdom and Australia): Breaking with competition and toxic relations**

'A sheltered workshop for gas bags' is what a previous head of department called universities. I must have missed the sheltered workshop part because my experience of university departments was the opposite of that. Margaret Thatcher's suspicion that not everyone was pulling their weight, and her slavish belief in metrics, led, in 1986, to the introduction of the first Research Assessment Exercise in the UK. My entry into academia occurred some five or six years later so it is likely that the gas bags were already being weighed and found wanting by the time I entered the fray. The history and subsequent development of how research and other academic output was and is measured can be found in the literature above, and elsewhere. But there is no question that the common purpose in academia is to bring in money through publication and research grants. Discovery, innovation, and academic excellence are also the purpose, but it's about money. Thatcher has had a long reach.

My various jobs—let's call it a career—have taken me to senior positions in both academia and government service. An academic department is, I am convinced, a unique work

environment. My experience in the five universities in three different countries in which I worked was characterised by an ongoing sense of not fitting in, of not understanding the undercurrents, of constantly feeling not part of the club: *not belonging*. Joining a department is like joining a family wedding or a wake halfway through. All the family members are in role, and there is a sense that something ominous is about to happen: a fight breaking out, perhaps, or old hurts being dragged up and played out. It's confusing for a new staff member who does not yet understand the undercurrents. Nor do they understand the jealousies and rivalries between all the relatives; how they started and why people are so exercised about what appear to be small matters to an outsider. Department meetings are characterised by meaningful looks across the room, notes passed to one's neighbour, raised eyebrows, sniggers, and even guffaws.

I speculate whether the personal attributes of a good academic may be also those that also make you an awful person to work with. To bring in money you must be better than the next person, and you need to be innovative; in other words, you must compete. It's a marvel how nasty people can be on the way up the ladder. A Faculty Dean described it as 'clever people thinking up clever ways of being horrible to each other'. It is a job that requires a tremendous amount of ego and opportunities to indulge in truly intellectual thought and open discussion I found to be strangely rare. Ego has no place in government service. And not to be too naive about it, government service, is public service and is based on a common purpose. However, no cloud comes without a diamond-dazzling lining, so towards what was to be the end of my academic career, I switched jobs and entered a fascinating public service role. Having a PhD seems to hold a lot of sway in contexts other than the academy! I'm glad I'm 'out' now. I will always operate as a supporter of those in higher education but from an emeritus position. Will I work in a university again, though? One never knows.

#### **Narrative 5 (Australia and New Zealand): Going it alone: The para-academic**

There are few things more liberating than working in an 'emeritus' or 'honorary' capacity. You have earned your badge and stripes, can work on passion projects and also support, as in my case, graduate students. You can be the 'old world' learner-focused and research-driven academic, and your energies are your own. You may no longer belong to a university, where 'belong' means 'have a reciprocal capital-based relationship with', but you can belong, finally, to yourself. The space of the independent is a rewarding one for those who made their mark, and those who have simply had enough.

The generations known as 'Gen X' and 'Gen Y' were those most hit by managerialist and neoliberalist reforms, and ultimately these were the majority of those whose positions were lost due to the endless restructures and redundancy rounds of the past decades (Andrew, 2020). COVID-19 afforded opportunities to thin the academic ranks further. Around me, colleagues were left without tenured affiliations

and consoled with honorary ones, which amount to online library access. They were asking whether they wanted another such position; many realised enough had been enough long ago, and thus went on to occupy independent academic spaces or to work in contexts of reinvention where they discovered at least fleetingly what workplaces had been like before the fall of the university.

At the centre of a group, I heard many stories. Some of the titles might be *Death by Administrivia* (and other 'Death by...' titles), *The Annual Crisis*, *Micromanagement Survivors Anonymous*, *The Sycophancy of the Neoliberal and Bullied to Death*. I carry these stories with me, but elect, for myself, to aim to make a difference. While I work at a distance with postgraduate learners, the pandemic made us all realise how distance, and Zoom and its proxies, were enablers of the less affiliated and more agentive life. The space of contract work afforded possibilities, both pedagogic and ecological. We realised the campus was, in fact, unnecessary, and had long since ceased to offer community to either graduate learners or their mentors.

We created a community afresh online, initially because we had to, but ultimately realised this saved time, stress, and horsepower. The pandemic catalysed two trends already underway: the fresh possibilities of online one-on-one, peer, and group supervision and the necessity of the unaffiliated academic. Fostering teams or communities of belonging for postgraduate learners testifies to the human need to align with professional or social groups as a motivation and support mechanism; but working independently from an institution suggests that for academics of my generation, there is more agency and authenticity in going it alone.

### **Narrative 6: (New Zealand) A new start with the same identity**

Just prior to the turn of the century, I had felt I found a place of belonging when I achieved a position at the university that I'd long wanted to be part of. I performed well within the increasing audit culture, but increasingly with less heart and more stress. Then there came a 'camel's back' moment. I have to say that there are many serious push factors from my work as a senior lecturer at a prominent university; after 20 years there, the gloss has well and truly worn off, and it has become a bit of a toxic work environment for me at least.

I will keep the backstory short: I spearheaded a protracted and bitter battle a couple of years ago to get casual teachers in our school made permanent, which I eventually won, with the excellent help of the union, but it took a toll on me. I later discovered the Vice Chancellor of my organisation was allied with the Atlas Think Tank, which cross-pollinated a particularly vile form of neoliberalism at that time and continues to threaten democracy itself today. At least I am departing, having made life better for 12 of my colleagues, so there is a grim satisfaction in that.

Although the push factors resulted from pain and disappointment, these were outweighed by new hope. Pull factors towards the new, and largely online educational

organisation—not a university—are a fresh start, a permanent, full-time position with a much better salary, seemingly lovely people, and the feeling of being valued rather than disposable, replaceable. So, I am moving from my core disciplinarity now towards related specialisms which have always been one strand of my career. I am also stepping out of the university sector. The air is fresher and cleaner, and I can breathe again.

### **Discussion**

The narratives testify to four things. First, the instrumentality and surveillance characterising the modern university prove major push factors to passionate educators; second, there is intellectual life after a university 'career', as Narrators 4 and 6 still call that sequence of random events. Third, a more ethical and authentic sense of academic identity is possible, too and it is here where the possible categories of public intellectual and para-academic apply, as Narratives 2, 3, 4, and 5 tell of ongoing multiple academic and emeritus roles beyond the hallowed walls. Narrative 2 aligns particularly with the ideal of the public intellectual. These are the extramural, extracurricular educationalists Withers and Wardrop (2014) described. Narratives 1 and 6 tell of moving out of toxic workplaces into places of new hope where the academic identity can be restored. Fourthly, the narratives unveil what Gill (2009) called hidden injuries and unmask the neoliberal university's failures in upholding collectives' and individuals' duties of care, especially in Narrative 1's story of management silencing those who speak out and trapping knowledge of change within an inner circle and Narrative Five's battle for integrity to improve conditions for others. Yet the stories also speak to individuals' critical resilience (Bottrell & Keating, 2019) as in the narratives' reversals of fortune, even if it brings what Narrative 6 calls a "grim satisfaction".

Whilst just one of the narratives mentions public intellectuals or para-academics (Narrative 5), there is a sense of "relational being" (Graham & Moir, 2022) in each narrative. That is, each individual nominates a new educational identity. Narrative One is a transition to a less stressful, academic position, while Narrative Two operates an independent academic existence beyond official ivory towers as a mentor, supervisor, researcher and activist. Narrator 1 is torn from their passion but worked down by the panopticon of managerialist discipline and punishment; Narrator 2 is convinced an authentic academic existence is impossible under neoliberalism. However, a rewarding and ethical intellectual life is possible, though corridor conversations are missed. The third narrator's key theme is the psychological impact of ingrained bullying, and how a neoliberalist human relations regime can turn a victim into a miscreant, themes seen, too, in Narratives 1 and 6. The survivor of bullying, even sheltered in a research institute, bears the scars. Trusting others and making collegial friends will always be a challenge.

Narrator 4 is not resistant to the capitalist, competitive thrust of the university, but wonders why they are environments that attract effective narcissists and sociopaths. They wonder about a connection between academic social climbing and



awfulness as a personality trait. They speak of not belonging to five universities across three countries due to institutional politics and ladder-climbing egos. Organisational histories that remain unresolved impact newcomers. On leaving, the joy of reading and writing for pleasure returns. Narrator 5 revels in finally being their own person, creating an academic identity that is fulfilling behind the restrictive parapets. They also cast themselves as a representative of a generation hit hardest by neoliberalism's tight fist and name the push-factors covered in the literature review. In a sense, they carry others' similar stories within them. The pandemic forged possibilities of online belonging to professional and postgraduate groups, giving voice to the value of the outsider. Narrator Six speaks of a transition from a stable role, through an organisational crisis, to a fresh role outside the university sector, but one which uses core expertise. She hints at the push factor of a Vice Chancellor whose work as an operative for the Atlas Network severely altered the direction of her university and many of its academics; her short vignette now seems like a window onto a terrifying emerging story (Hamilton, 2024).

A troubling theme of the feeling of being bullied runs through five of the six narratives. Narrative 4 called it "clever people thinking up clever ways of being horrible to each other". Narrator 2 calls it "sanctioned"; the neoliberal university condones such behaviour by perpetuating justice imbalances and privileging hegemonic voices. Even with union activism, there is a sense that neoliberalism will be victorious. A metaphor of 'battle' (Narrative 6) rings throughout. Also thematic is the idea that the university is a petri dish for sycophantic neoliberals, anxious for a place on the ladder at any cost. Narrative 6 hints at a powerful lobby group influencing the expression of this neoliberalism in higher education, and their infiltration by vested power and interests. A theme of academic hard work or success not bringing joy prevails. A refreshing theme of finding breathing space is present, too (Narrators 1, 3 and 6), as is the idea of work as a public intellectual being ethical, with professional probity a function of the profession itself, not the organisation (Narrators 3 and 5). Significantly, most narrators note the connection between inferred status and affiliation. For Narrators 1, 5 and 6, the sense of still being able was confirmed with a new offer. The archaic notion of 'affiliation' is still required in technocratic systems, circulated in conference communities and normalised by publishers who often require an affiliation as part of the submission process as if there is no possibility of identifying as an independent researcher for purely bureaucratic reasons. Non-affiliated independents exist. The symbolic violence of being made redundant, discarded or moved-on wounds us psychologically (Andrew, 2020). It "takes a toll" (Narrative 5). Not being affiliated is an identity option that affords agency.

## Conclusions

As the 2020s run on, an emancipatory counter-narrative (Goodall, 2010) where the academic is agential continues to evolve, spurred by such critiques as that of Neary and Wynn (2016). Perhaps also Utopian, this sub-genre is activist, malleable and adaptive, and my study has pivoted on appreciating such things as critical resilience, para-

academic work, and the joys of working beyond affiliation; of *not* belonging to a badged institution. Not belonging is here an agentive state that refigures power, opportunity, and identity in authentic ways beyond toxic universities and their zombiedom (Smyth, 2017). It resists the logically ludicrous notion that one single monolithic organisation defines individuals and opens the possibility of operating independently under multiple banners or realigning skills to other professional endeavours. It is concomitant with notions of liquid modernity with its information overload and super speedy change (Stein, 2021) where ex-academics, para-academics, public intellectuals, and those in emeritus capacities contribute broadly to global citizenship education and even cosmopolitanism in the sense of non-affiliation. Should their traumas persist, they may seek "to develop and disseminate a revised set of shared educational ideals, values, and modes of meaning making and social change" pivoting on "democracy, inclusion, and shared humanity" (Stein, 2021, pp. 483, 486).

Thriving outside the academy as ethical, creative non-conformists is increasingly an option, even if it is a plan B (Barcan, 2017). This "paraversity" of "para-academics" offers the ability to do good academic work outside institutions as creators, experts, contractors, or consultants, operating unseen in plain sight (Rolfe, 2014; Withers & Wardrop, 2014). Withers and Wardrop (2014) write that such extracurricular educationalists "carve out opportunities to inhabit spaces that appear off limits under the terms of the contemporary academy... so thoroughly 'occupied' by marketization" (p. 7). The universities are 'occupied' as by an enemy wartime force and even by "operatives" for corporate think tanks (Hamilton, 2024). Yet, it is possible to occupy a new place for articulating and reclaiming the value and integrity of practical and collective work of knowledge and resistance. Clearly, the demand for 'affiliation' needs to disappear from many a technocratic apparatus such as conference enrolment forms; without independent thought beyond group/think/tanks, the academic is rudderless. And without being allowed integrity, the designation 'academic' becomes a falsehood, a non-entity.

Thriving is not only about an ongoing intellectual life, but also about well-being. Acton and Glasgow (2015) argue that only contexts that "provide possibilities for action, agency and autonomy" can be seen as supportive and remedial of pressures (p. 107). Their study is particularly damning of the death by administrative workloads, "the metric" (Story 4) and "audit culture" (Story 6) that comes with intensification in neoliberal universities. Narrator Two could survive only by flying under the radar in a geographically-other research institute beyond the panopticon. The lived experience of the six narrators shows that no amount of passion, talent, and success can counterbalance administrivia. All stories testify to the academic incivility and bully culture that Twale and DeLuna (2008) identified. There is a lack of morale, that in Brion's (2015) conception, comes when belonging doesn't exist. However, thriving offers the possibility of not-belonging as an option.

Thriving seems, too, to be about different forms of belonging, but an authentic sense of non-hegemonic belonging, not that of organisational propaganda and

agitprop orientations. Scholarly belonging differs from organisational belonging: it involves mentoring others, building capacity and capability through the exercise of one's experience; it involves a common purpose, as in our narrators' stories of finding new places, potentially in spaces of "democracy, inclusion, and shared humanity" (Stein, 2021, p. 486). It is simultaneously a selfless and a human need, evolves organically, and cannot be enforced top-down by technocracy. Most of all, the thriving that happens in these stories happens independent of affiliation to the modern university and enables an authentic academic identity with integrity.

### Further research

Lewin (2023) reported on the post academia careers of five academics during 'the great resignation'. A phenomenon of mass exodus by choice or redundancy is caused more by push factors than pull factors: there are more stories of escape from toxicity than finding a new workplace for fresh belonging. A sense that teaching and learning are not what they should be due to a loss of educator agency pervades. A 2019 article in *The Chronicle of Higher Education* offered the evocative title: "'This was a hell not unlike anything Dante conjured': Readers share their stories of fraught academic careers". Clearly, there is more to tell beyond my, Lewin's (2023) or Barcan's (2013, 2017) scope about the phenomenon of 'the great resignation', specifically issues of educator agency, organisational policy and even curriculum content that need exploring with a view to resulting in positive change. The unlikelihood of positive change, at least in the present, is arguably because of the cockroach nature of the neoliberal reported here and elsewhere (Ball, 2015; Fleming, 2021; Andrew, 2020, 2023, 2024). Hence, possibilities for innovation are limited without change due to the all-pervading ideological palsy of neoliberalism (Hil, et al., 2022). Further, the ethics of care that I owe to narrators requires me to stay with impressions, perceptions and experiences and prohibits me from reporting details that may damage institutions or reflect poorly on surviving educators teaching their changed/sabotaged curricula.

I have already written of how neoliberalist process has itself no room for care in assessment processes even in the post-COVID world (Andrew, 2024) and of how educators remaining after 'the great resignation' may be *homo economicus* clones, tow-the-line zombies or conscious-riven nervous wrecks (Andrew, 2023). It is also important to ask the question, 'Whatever happened to the exit interview?' and the concomitant enquiries, 'Did exit interviews ever do anyone any good?' and 'Did organisational change ever result from exit interview data?' The truth is that, for universities, except in the case of some who stay in emeritus or honorary capacities, once you're gone, you're gone. The reality is that exit interviews, once an aspect of process and duty of care, seldom if ever occur anymore.

One last question that needs to be asked, and which will take considerable courage, considers the possible impact of the loss of ideological outsiders from democratic processes of teaching and learning. The loss of these educators represents a massive loss for present and future generations.

This question lies in the domain of Giroux's 2019 horrifying work on higher education in a populist universe: *Terror of the unforeseen*.

### References

- Acton, R., & Glasgow, P. (2015). Teacher wellbeing in neoliberal contexts: A review of the literature. *Australian Journal of Teacher Education*, 40(8), 99-114. <http://dx.doi.org/10.14221/ajte.2015v40n8.6>
- Andrew, M. (2020). Behind voluntary redundancy in universities: The stories behind the story. *Australian Universities' Review*, 62(2), 14-24. <https://search.informit.org/doi/10.3316/informit.398447476011801>
- Andrew, M. (2023). Come to the cabaret: Voices from the modern university. *Journal of Applied Learning and Teaching*, 6(2), 17-27. <https://doi.org/10.37074/jalt.2023.6.2.19>
- Andrew, M. (2024). "Just get them over the line": Neoliberalism and the execution of 'excellence'. *Journal of Applied Learning and Teaching*, 7(1), 1-11. <https://doi.org/10.37074/jalt.2024.7.1.25>
- Antonsich, M. (2010). Searching for belonging – an analytical framework. *Geography Compass*, 4(6), 644-659. <https://doi.org/10.1111/j.1749-8198.2009.00317.x>
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215-228. <https://doi.org/10.1080/0268093022000043065>
- Ball, S. J. (2012). Performativity, commodification and commitment: An I-spy guide to the neoliberal university. *British Journal of Educational Studies*, 60(1), 17-28. <https://doi.org/10.1080/00071005.2011.650940>
- Ball, S. J. (2015). Living the neo-liberal university. *European Journal of Education*, 50(3), 258-60. <https://www.jstor.org/stable/26609276>
- Barcan, R. (2013). *Academic life and labour in the new university: Hope and other choices*. Ashgate. [https://www.researchgate.net/publication/286403112\\_Academic\\_life\\_and\\_labour\\_in\\_the\\_new\\_university\\_Hope\\_and\\_other\\_choices](https://www.researchgate.net/publication/286403112_Academic_life_and_labour_in_the_new_university_Hope_and_other_choices)
- Barcan, R. (2017). Universities need to plan for a dark future if academics prefer their own Plan B. *Times Higher Education*. <https://world.edu/universities-need-plan-dark-future-academics-prefer-plan-b/>
- Barker, J. (2017). Ninjas, zombies and nervous wrecks? Academics in the neoliberal world of physical education and sport pedagogy. *Sport, Education and Society*, 22(1), 87-104. <http://dx.doi.org/10.1080/13573322.2016.1195360>
- Berg, L. D., Huijbens, E. H., & Larson, H. G. (2016). Producing anxiety in the neoliberal university. *The Canadian Geographer/Le Geographe Canadien*, 60(2), 168-180. <https://doi:10.1111/cag.12261>

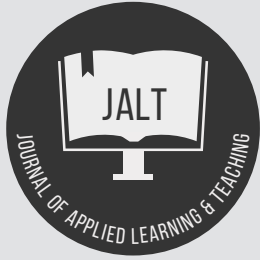
- Berryman, M., & Eley, E. (2019). *Student belonging: Critical relationships and responsibilities*. <https://researchcommons.waikato.ac.nz/handle>
- Bottrell, D., & Keating, M. (2019). Academic wellbeing under rampant managerialism: From neoliberal to critical resilience. In D. Bottrell & C. Manathunga (Eds.), *Resisting neoliberalism in higher education volume I* (pp. 157-178). Palgrave Critical University Studies. [https://doi.org/10.1007/978-3-319-95942-9\\_8](https://doi.org/10.1007/978-3-319-95942-9_8)
- Bottrell, D., & Manathunga, C. (Eds.) (2019). *Resisting neoliberalism in higher education volumes I and II*. Palgrave Critical University Studies. <http://dx.doi.org/10.1007/978-3-319-95834-7>
- Brion, S. A. (2015). *Teacher morale: A dissertation in educational leadership* [Doctoral dissertation, The Pennsylvania State University]. [https://etda.libraries.psu.edu/files/final\\_submissions/10546](https://etda.libraries.psu.edu/files/final_submissions/10546)
- Chronicle of Higher Education. (2019). 'This was a hell not unlike anything Dante conjured': Readers share their stories of fraught academic careers. <https://www.chronicle.com/article/this-was-a-hell-not-unlike-anything-dante-conjured-readers-share-their-stories-of-fraught-academic-careers/>
- Crawford, N. L., Emery, S. G., Allen, P., & Baird, A. (2022). I probably have a closer relationship with my internet provider: Experiences of belonging (or not) among mature-aged regional and remote university students. *Journal of University Teaching & Learning Practice*, 19(4). <https://ro.uow.edu.au/jutlp/vol19/iss4/11>
- Dews, C. L. B., & Law, C. L. (Eds.). (1995). *This fine place so far from home: Voices of academics from the working class*. Temple University Press. <https://www.jstor.org/stable/j.ctt14bswxr>
- Evans, L., & Nixon, J. (2015). *Academic identities in higher education: The changing European landscape*. Bloomsbury. [https://www.researchgate.net/publication/276273686\\_Academic\\_Identities\\_in\\_Higher\\_Education\\_The\\_Changing\\_European\\_Landscape](https://www.researchgate.net/publication/276273686_Academic_Identities_in_Higher_Education_The_Changing_European_Landscape)
- Flaherty, C. (2022, July 4). Calling it quits. *Inside Higher Ed*. (July, 4). <https://www.insidehighered.com/news/2022/07/05/professors-are-leaving-academe-during-great-resignation>
- Fleming, P. (2021). *Dark academia: How universities die*. Pluto Books. <https://doi.org/10.2307/j.ctv1n9dkhv>
- Gill, R. (2009). Breaking the silence: The hidden injuries of neo-liberal academia. In R. Flood & R. Gill (Eds.), *Secrecy and silence in the research process: Feminist reflections* (pp. 228–244). Routledge. <http://dx.doi.org/10.1515/fs-2016-0105>
- Ginsberg, B. (2011). *The fall of the faculty*. Oxford University Press. <https://doi.org/10.1093/oso/9780199782444.001.0001>
- Giroux, H. A. (2014). Public intellectuals against the neoliberal university. In N. K. Denzin & M. D. Giardina (Eds.), *Qualitative inquiry outside the academy* (pp. 35–60). Left Coast Press. <http://dx.doi.org/10.4324/9781315421339-1>
- Giroux, H. A. (2017). Neoliberalism's war against higher education and the role of public intellectuals. In M. Izak, M. Kostera, & M. Zawadzki (Eds.), *The future of university education* (pp. 185-206). Palgrave. [https://doi.org/10.1007/978-3-319-46894-5\\_10](https://doi.org/10.1007/978-3-319-46894-5_10)
- Giroux, H. A. (2019). *Terror of the unforeseen*. Larb.
- Goodall, H. L., Jr. (2010). *Counter-narrative: How progressive academics can challenge extremist and promote social justice*. Left Coast Press. <https://doi.org/10.4324/9781315431499>
- Graham, C. W., & Moir, Z. (2022). Belonging to the university or being in the world: From belonging to relational being. *Journal of University Teaching & Learning Practice*, 19(4). <https://ro.uow.edu.au/jutlp/vol19/iss4/04>
- Hamilton, L. (2024). *New Zealand is under siege by the Atlas Network*. The Australian Independent Media Network. [https://theaimn.com/new-zealand-is-under-siege-by-the-atlas-network/?fbclid=IwAR24sqmt\\_8exGV8hEDXpqKB7arhRksgokYxe7mhNgRniBTkm5WtmtB35qQ](https://theaimn.com/new-zealand-is-under-siege-by-the-atlas-network/?fbclid=IwAR24sqmt_8exGV8hEDXpqKB7arhRksgokYxe7mhNgRniBTkm5WtmtB35qQ)
- Healey, K., & Stroman, C. (2021). *Structures for belonging: A synthesis of research on belonging-supportive learning environments*. Student Experience Research Network. <https://studentexperiencenetwork.org/wp-content/uploads/2021/03/Structures-for-Belonging.pdf>
- Hil, R., Pelizzon, A., & Baum, F. (2022). It's time: The reform of Australian private universities. *Social Alternatives*, 41(1), 3-4. <https://socialalternatives.com/issue/its-time-the-re-form-of-australian-public-universities/>
- Hung, D., & Der-Thang, C. (2001). Situated cognition, Vygotskian thought and learning from the communities of practice perspective: Implications for design of web-based e-learning. *Educational Media International*, 38(1), 3-12. <http://dx.doi.org/10.1080/09523980121818>
- Kandemir, A., & Budd, R. (2018). Using vignettes to explore reality and values with young people. *Forum Qualitative Sozialforschung*, 19(2), 1-23. [https://www.researchgate.net/publication/346492969\\_Using\\_Vignettes\\_to\\_Explore\\_Reality\\_and\\_Values\\_With\\_Young\\_People](https://www.researchgate.net/publication/346492969_Using_Vignettes_to_Explore_Reality_and_Values_With_Young_People)
- Kahu E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research and Development*, 37(1), 58-71. <https://doi.org/10.1080/07294360.2017.1344197>
- Koukouraki, K. (2020). Supporting students in developing critical global citizenship: Examples from the English for Academic Purposes (EAP) classrooms. *Journal of Applied Learning & Teaching*, 3(2), 107-115. <https://doi.org/10.37074/jalt.2020.3.2.16>
- Lähdesmäki, T., Saresma, T., Hiltunen, K., Jäntti, S., Säskilähti, N., Vallius, A., & Ahvenjärvi, K. (2016). Fluidity and flexibility

- of 'belonging': Uses of the concept in contemporary research. *Acta Sociologica*, 59(3), 233-247. <https://doi.org/10.1177/0001699316633099>
- Lewin, V. (2023, November 6). *How five researchers fared after 'their great resignation' from academia*. Nature. <https://www.nature.com/articles/d41586-023-03484-7?fbclid=IwAR14-bdMR7tdshJwtzikst3ZzUPZpcyeVm3khXMEVhm4hlf-TQI0MOHuUSM>
- Loads, D., Peseta, T., Smith, J., & Rattray, J. (2016). Identity work in the contemporary university: Exploring an uneasy profession. *Sense*. <http://dx.doi.org/10.1007/978-94-6300-310-0>
- Mulrooney, H. M., & Kelly, A. F. (2020). COVID-19 and the move to online teaching: Impact on perceptions of belonging in staff and students in a UK widening participation university. *Journal of Applied Learning & Teaching*, 3(2), 17-30. <https://doi.org/10.37074/jalt.2020.3.2.15>
- Neary, M., & Hagyard, A. (2010). Pedagogy of excess: An alternative political economy of student life. In M. Molesworth, R. Scullion, & E. Nixon (Eds.), *The marketisation of higher education and the student as consumer* (pp. 209–224). Routledge. <https://api.core.ac.uk/oai/oai:eprints.lincoln.ac.uk:4187>
- Neary, M., & Wynn, J. (2016). Against academic identity. *Higher Education Research & Development*, 35(2), 409-412. <https://doi.org/10.1080/07294360.2015.1094201>
- Owan, V. J., Agama, V. U., Odey, J. O., & Idika, D. O. (2024). Metrics in research impact assessment and grant funding: Insights from researchers in the "Reviewer 2 Must Be Stopped!" Facebook group. *Journal of Applied Learning & Teaching*, 7(1), 1-12. <https://doi.org/10.37074/jalt.2024.7.1.19>
- Polkinghorne, D. E. (1997). Reporting qualitative research as practice. In W. G. Tierney & Y. S. Lincoln (Eds.), *Representation and the text: Reframing the narrative voice* (pp. 3-21). State University of New York Press. <https://www.jstor.org/stable/3196030>
- Poulos, C. (2017). Autoethnographic reflections on the neoliberal academy: Stories of resistance, resilience, and remembrance. *Cultural Studies ↔ Critical Methodologies*, 17(4), 307. <https://doi.org/10.1177/1532708617706121>
- Press, N., Andrew, M. B., Percy, A., & Pollard, V. A. (2022). Pedagogies of belonging in an anxious world: A collaborative autoethnography of four practitioners. *Journal of University Teaching & Learning Practice*, 19(4). <https://ro.uow.edu.au/jutlp/vol19/iss4/01>
- Rolfe, G. (2014). We are all para-academics now. In A. Wardrop & D. Withers (Eds.), *The para-academic handbook: A toolkit for making-learning-creating-acting* (pp. 1-5). HammerOn Press. <http://doi.org/10.1386/9780956450753>
- Shore, C. (2008). Audit culture and illiberal governance: Universities and the politics of accountability. *Anthropological Theory*, 8, 278–298. <https://doi.org/10.1177/1463499608093815>
- Shore, C. (2010). Beyond the multiversity: Neoliberalism and the rise of the schizophrenic university. *Social Anthropology*, 18(1), 15-29. <http://dx.doi.org/10.1111/j.1469-8676.2009.00094.x>
- Slaughter, S., & Leslie, L. (1999). *Academic capitalism: Politics, policies and the entrepreneurial university*. Johns Hopkins University Press. <https://doi.org/10.5860/choice.35-3451>
- Smyth, J. (2017). *The toxic university: Zombie leadership, academic rock stars and neoliberal ideology*. Palgrave Macmillan.
- Sparkes A. C. (2007). Embodiment, academics, and the audit culture: A story seeking consideration. *Qualitative Research*, 7(4), 521-550. <https://doi.org/10.1177/1468794107082306>
- Stafford, V. (2022). Successful collaboration in online learning through skills and community building: A women in leadership MBA subject case study. *Journal of Applied Learning & Teaching*, 5(2), 160-168. <https://doi.org/10.37074/jalt.2022.5.2.ss4>
- Stein, S. (2021). Reimagining global citizenship education for a volatile, uncertain, complex, and ambiguous (VUCA) world. *Globalisation, Societies and Education* 19(4), 482-495, <https://doi.org/10.1080/14767724.2021.1904212>
- Taylor, P. (2013). Putting theory to work – a.k.a. 'if you don't like academia, why don't you leave?' *Ephemera: Theory & Politics in Organization*, 13(4), 851–860. <http://www.ephemerajournal.org>
- The professor is out. (2023). *Home* [Facebook page]. Facebook. <https://www.facebook.com/groups/professorisout/>
- Tight, M. (2000). *Academic work and life: What it is to be an academic, and how this is changing*. Elsevier. [https://doi.org/10.1016/S1479-3628\(2000\)1](https://doi.org/10.1016/S1479-3628(2000)1)
- Tight, M. (2010). The golden age of academe: Myth of memory? *British Journal of Educational Studies*, 58(1), 105-116. <https://doi.org/10.1080/00071000903516502>
- Tregear, P., Guthrie, J., Lake, S., Lucas, A., O'Connor, J., Pellzozon, A., & Vodeb, B. (2022). 'Enough to make you sick!' Pathological characteristics of the Australian academic workplace. *Social Alternatives*, 41(1), 44-51. [https://socialalternatives.com/wp-content/uploads/2022/04/TREGEAR-et-al-SA\\_41\\_1.pdf](https://socialalternatives.com/wp-content/uploads/2022/04/TREGEAR-et-al-SA_41_1.pdf)
- Twale, D. J., & DeLuna, B. M. (2008). *Faculty incivility: The rise of the academic bully culture and what to do about it*. Jossey-Bass. <https://doi.org/10.1353/rhe.0.0133>
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press. <https://psycnet.apa.org/doi/10.1017/CBO9780511803932>
- Withers, D., & Wardrop, A. (2014). Reclaiming what has

been devastated. In A. Wardrop & D. Withers (Eds.), *The para-academic handbook: A toolkit for making-learning-creating-acting* (pp. 6-13). HammerOn Press. <https://doi.org/10.1386/9780956450753>

Yuval-Davies, N. (2006). Belonging and politics of belonging. *Patterns of Prejudice*, 40(3), 197-214. <https://doi.org/10.1080/00313220600769331>

Copyright: © 2024. Martin Benedict Andrew. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Neoliberalism and the violation of students' rights: The case of English language education

---

Ehsan Nikpouya<sup>A</sup>

A

*Hakim Sabzevari University, Sabzevar, Iran*

Gholamreza Zareian<sup>B</sup>

B

*Hakim Sabzevari University, Sabzevar, Iran*

---

### Keywords

Education;  
ELT (English Language Teaching);  
neoliberalism.

---

### Abstract

This theoretical study explores the impact of neoliberalism on education, specifically focusing on its influence in the areas of general education and English language teaching. Neoliberalism, a complex concept with varied interpretations, advocates for free markets, privatisation, and individual entrepreneurship as a means to enhance human well-being. In the context of education, neoliberal ideologies have led to the commodification of education, the transfer of educational responsibility from the government to individual households, and the emphasis on standardised testing and accountability measures. The impact of neoliberalism is particularly pronounced in the English language teaching sector, where English is viewed as a valuable economic asset and a tool for individual advancement in the market-driven economy. Despite the growing awareness of neoliberalist influences, there is a need for further academic exploration and critical inquiry to understand its pervasive effects on education systems and learners. Regarding the matter, this study discusses the influence of neoliberal ideologies on textbooks and materials, which often promote themes of employability, consumerism, and entrepreneurship.

---

### Correspondence

[g.zareian@hsu.ac.ir](mailto:g.zareian@hsu.ac.ir) <sup>B</sup>

---

### Article Info

Received 5 September 2023

Received in revised form 26 December 2023

Accepted 6 January 2024

Available online 8 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.3>

## Introduction

Neoliberalism advocates for free markets, privatisation, and individual entrepreneurship as means to enhance human well-being. In the context of education, neoliberal ideologies have led to the commodification of education, the transfer of educational responsibility to households, and the emphasis on standardised testing and accountability. This theoretical study utilises different resources to examine the global infiltration of neoliberalism in educational systems. The study also discusses the influence of neoliberal ideologies on English Language Teaching (ELT) textbooks and materials, which often promote themes of employability and consumerism.

In terms of theory, the study delves into the multifaceted concept of neoliberalism. It acknowledges the challenge of providing a concise and universally agreed-upon definition due to the polarising responses and subjective perspectives surrounding it. Drawing on the works of Harvey (2005), Giroux (2008), Plehwe (2009), Brown (2015), and Andrew (2023), the study explores neoliberalism as a theoretical framework advocating for individual entrepreneurial freedoms and promoting the financialisation of various aspects of society. The study then examines the manifestation of neoliberalism in education, highlighting the shift in educational philosophy towards individualism and competitive business models. It discusses the transfer of educational responsibility from the government to households, the prevalence of standardised testing, and the privatisation of public goods such as education. Drawing on the works of researchers like Ball (2003), Savage (2017), and Hastings (2019), this study explores the global influence of neoliberal governance on educational institutions, policymakers, and curriculum designers.

Furthermore, this paper examines the impact of neoliberalism, specifically in the context of English language teaching. It explores the historical association of English with colonial powers and the ideological debates surrounding ELT. This study draws on the works of Bernstein et al. (2015), Babaii and Sheikhi (2017), Bacon and Kim (2018), Bori and Canale (2022), and Farsani and Rahimi (2023) to discuss how neoliberal ideologies conceptualise English as a marketable skill and viewing individuals primarily as economic assets in the globalised world. It also addresses the influence of neoliberalism on ELT textbooks and materials, emphasising themes of employability and consumerism. The significance of the matter at hand is undeniable. However, the current body of research on this topic is insufficient, with most studies failing to directly address the issue. As a result, we have undertaken a comprehensive research endeavour to provide a well-structured and informative analysis of this matter. Our aim is to assist English language researchers and teachers in gaining a clear and concise understanding of this topic.

Thus, the present study addresses the following research questions:

1. How is neoliberalism conceptualised by different researchers?

2. How is neoliberalism manifested in mainstream education?
3. What are the traces of neoliberalism in English language education?

## Theoretical considerations: Neoliberalism

The task of defining 'neoliberalism' in 2023 is notably challenging due to the presence of polarising responses, both in favour and against the concept, as well as the varied forms and adaptations it has taken (Andrew, 2023). As Plehwe (2009) puts it, neoliberalism encompasses a range of intricate and multifaceted perspectives, lacking a concise and universally agreed-upon definition within the realm of political philosophy. Since the 1970s, there has been a pervasive inclination towards neoliberalism evident in both political-economic practices and ideological perspectives around the world. The prevalence of deregulation, privatisation, and the state's retreat from various domains of social provision has been a frequent occurrence. Vandrick (2014) asserts that these policies contribute to the perpetuation and amplification of inequities in social class statuses.

According to Harvey (2005, p. 33), neoliberalism can be characterised as the "financialisation of everything". He states that neoliberalism primarily represents a theoretical framework concerning political economic practices, advocating for the promotion of individual entrepreneurial freedoms and abilities as essential means to enhance human well-being. This ideology emphasises the establishment of robust private property rights, unrestricted markets, and unhindered international trade as key components of the institutional framework. The phenomenon of neoliberalisation has resulted in extensive instances of radical rather than transformative change, commonly referred to as "creative destruction" (Harvey, 2005, p. 3). These changes encompass not only the dismantling of existing institutional frameworks and sources of authority (sometimes even undermining conventional forms of governmental control), but also the reconfiguration of labour divisions, social interactions, welfare arrangements, utilisation of mixed technological approaches, cultural norms, cognitive processes, reproductive behaviours, emotional connections to the environment, and deeply ingrained personal inclinations (Harvey, 2005). Giroux (2008, p. 9) characterised neoliberalism as "a broad-based rhetorical and cultural movement designed to obliterate public concerns and liquidate the welfare state".

Brown (2015) also states that neoliberalism is commonly recognised as implementing a wide range of economic policies that align with its core principle of supporting free markets. These policies encompass the deregulation of industries and capital flows, a significant reduction in social welfare provisions and protections for vulnerable individuals, and the privatisation and outsourcing of public goods such as education, parks, postal services, roads, social welfare programs, prisons, and militaries. Additionally, neoliberalism involves replacing progressive tax and tariff systems with regressive ones and abandoning wealth redistribution as an economic or socio-political

policy (Brown, 2015). It also involves the transformation of every human need or desire into profitable endeavours, spanning from activities like college admissions preparation and human organ transplants to baby adoptions, pollution rights, and even trivial matters like avoiding queues or securing more legroom on airplanes. In addition, there has been a recent trend towards financialisation, where finance capital assumes greater dominance over productive capital in the dynamics of the economy and everyday life.

Michel Foucault, as a key figure also has insights on neoliberalism in his famous book, "The birth of biopolitics: Lectures at the Collège de France," which was published between 1978 and 1979. Based on the interpretation of Read (2009) to Foucault, neoliberalism represents a novel form of "governmentality," denoting a particular approach or mindset through which individuals are both governed and self-governed (p. 29). Neoliberalism effectively alters the language and concepts through which people situate themselves within society, shifting away from notions of "rights and laws" and towards considerations of "interests, investments, and competition". Paradoxically, as a mode of governance, neoliberalism appears to govern while seemingly not governing directly; in order to operate, its subjects require considerable autonomy to make choices amid competing strategies (Read, 2009, p. 29).

## Neoliberalism in education

Hastings (2019) asserts that neoliberalism portrays individuals as rational actors driven by self-interest, primarily tasked with maximising their personal well-being. He accounts for what the status of education is with regard to neo-liberalism in five ways as the following:

1. Neoliberal ideology portrays education as a financial venture, wherein education is regarded primarily as an economic investment.
2. The burden of educational provision is transferred from the government to individual households, thereby placing families in charge of the responsibility for education.
3. Neoliberal education reform relies on standardised testing to assess and quantify school worth, enabling informed decisions on investments and enrolment for policymakers, parents, and students.
4. Neoliberalism views education as a technical rather than a political issue, emphasising career and college readiness. Test scores are used to gauge preparedness, with schools prioritising research on methods to enhance performance.
5. Investors in the privatisation of public education aim to profit by providing contracting, testing, tutoring, school management, and non-educational services like marketing to charter schools.

Although the manifestations of neoliberalism have exhibited variation across different regions, the whole global education system has been significantly influenced by neoliberal modes of governance (Savage, 2017). For example, Australia has witnessed substantial restructuring of its government schooling systems and curricula as a result of neoliberalism. This trend persists as reforms increasingly align with market-driven principles. Over the past twenty years, Australian governments have extensively adopted neoliberal policies from the United Kingdom and the United States, disregarding extensive research that illuminates the adverse consequences of these reforms on schools, educators, and students. Stephen Ball (2003), a prominent theorist in the field of education policy, posits that the proliferation of educational reforms shaped by neoliberal ideologies has traversed geographical boundaries, resembling what he terms a "policy epidemic" (p. 215). Prominent educational enterprises, such as Pearson, have significantly capitalised on these shifts by generating educational materials and securing profitable governmental agreements for the provision of fundamental educational services (Hogan, 2016).

The infiltration of neoliberalism into European educational policies faced significant opposition and generated both vocal critics and passionate supporters (Hakala et al., 2015). The ensuing tensions between researchers, teachers, and public officials came to the forefront during the 1980s and 1990s (Husén, 1994, 1996). Despite facing criticism, neoliberal principles were adopted more swiftly in Nordic educational systems compared to those in Continental Europe (Rinne, 2004). Finland, like other countries, began questioning the benefits of a compulsory educational system and sought to tailor schools to individual needs. Schools were recognised as integral components of the national economy, with their output seen as "human capital" (Ahonen, 2002, pp. 177–180; Kalalahti & Varjo, 2012, p. 48). By the mid-1990s, Finnish educational policy had embraced neoliberal values such as efficiency, competitiveness, and a focus on outcomes. Here it should be noted that the phenomenon of neoliberalisation in education is not exclusive to any particular nation but rather represents a widespread global trend that manifests variably at the local level (West, 2019).

Savage (2017) claims that the rise of neoliberalism has brought about a fundamental metamorphosis in the operational dynamics of educational institutions, the approaches employed by parents to navigate through school systems, the strategies pursued by policymakers to govern schools, the decision-making processes undertaken by curriculum designers regarding the knowledge and skills imperative for young individuals, as well as the conduct of school principals and teachers within educational establishments. The influence of neoliberalism on educational governance and schooling remains persistent, with limited indications of diminishing. Contrary to being perceived as a fading orthodoxy, neoliberalism maintains its energy, exerting radical effects on various dimensions of schooling, curriculum development, and the experiences of educators and students worldwide.



Popenici et al. (2023) argue that even universities are facing challenges from neoliberal ideology and a strong emphasis on transforming education into a commercial enterprise, with a primary focus on financial gains and market forces.

## Neoliberalism in English language education

The impact of neoliberalism on the education sector has been profound, resulting in a significant change in educational ideology. This change involves a shift away from valuing social cooperation and towards prioritising individualism and competitive business models (Block et al., 2012). English language teaching has not remained untouched by this global trend. Instead, due to its integral role in the process of globalisation and the historical association of English with colonial powers in many parts of the world, ELT has consistently grappled with ideological debates and a diverse range of influences (Babaii & Sheikhi, 2017).

According to Bernstein et al. (2015), the ideology of neoliberalism has conceptualised the English language as a marketable and technologically driven ability, while also viewing individuals primarily as valuable economic assets. Neoliberal ideologies and implementations within the realm of education sustain a perspective that regards language as an isolated proficiency, devoid of its contextual and societal connotations and ramifications (Bori & Canale, 2022). Textbooks are not neutral. The purpose of textbooks is to exert influence on learners. Textbooks serve as significant tools within the realm of education and socialisation, with the objective of preparing individuals to become responsible members of established societies. This preparation is achieved through the transmission of prevailing macro-level ideologies, essentially moulding individuals in a process commonly referred to as 'social reproduction' (Bourdieu & Passeron, 1977). Despite the considerable body of research on textbook ideology, there exists a dearth of scholarly investigation into the phenomenon of neoliberalism within the domain of ELT. Specifically, the manner in which neoliberalism establishes its pervasive influence within the English teaching industry and enlists individuals through educational pathways remains an underexplored area warranting further academic inquiry (Babaii & Sheikhi, 2017).

Under the neoliberal paradigm, English is perceived as a highly valuable yet impartial form of capital that can be accessed by individuals through diligent investment of effort and time (Heller & Duchêne 2012; Heller, 2003, 2010; Park, 2016; Urciuoli, 2008). Consequently, individuals are often held responsible for undertaking the necessary investment in acquiring a socially-expected level of English proficiency to thrive in an intensely competitive neoliberal market. Within this context, the ideology of English as a neutral skill and a pristine medium of potentiality has engendered a remunerative English language industry in South Korea (Park, 2016). The teaching of languages, particularly English as a foreign language, has transformed into a lucrative industry for the English-language sector. (Philipson, 1992; Barnawi, 2017). The acquisition of a second language is commonly framed within discursive frameworks as an ongoing endeavour aimed at continual personal advancement (Shin, 2016). Moreover, it is often

regarded as a strategic investment in one's own capabilities to enhance competitiveness within the realms of education and employment. Horiguchi et al. (2015) contend that neoliberal ideologies prioritise the notion that individuals bear the responsibility of attaining crucial information and skills, such as communication or language proficiency, which are regarded as significant matters within the contemporary knowledge-based economy. Smith (2022) argues that the implementation of a market-oriented approach to English language policy can be seen as an exclusive form of repression. This approach amalgamates forceful measures and consent, utilising strategies that legitimise and sustain the prevailing social, political, economic, and linguistic norms within a given context.

In their research findings, Bacon and Kim (2018) discovered that the participants' discussions on the use of English revealed a minimal connection between language acquisition and authentic interpersonal communication. Instead, their focus shifted towards leveraging their English test scores and perceived proficiency for social, educational, and peer advantages. Moreover, they recognised the practical necessity of English proficiency as a prerequisite for accessing educational and employment opportunities. The American Council for the Teaching of Foreign Languages' publication, the Standards for Foreign Language Learning: Preparing for the 21st Century (1996), and the Common European Framework of Reference (CEFR) (2001) established by the Council of Europe have played a substantial role in shaping this perspective. These documents provide standardised criteria for assessing language proficiency. However, according to Canale (2022), relying on metrics, descriptors, and evaluation parameters may create a sense of detachment between learners and the sociopolitical contexts in which language learning takes place. The standardised procedures employed by these documents have implications beyond assessing and evaluating language skills. For example, numerous European countries and even non-European nations utilise the CEFR for Languages to develop language courses and select teaching materials (Byram & Parmenter, 2012; Savski, 2020), which often encourages the adoption of a "teaching to the test" approach. Additionally, the CEFR has become a marketable brand, leading to the sale of language products like textbooks and exams.

In a study carried out by Abdollahzadeh and Baniasad (2010), contrasting attitudes towards the English language and its significance in relation to English society, individuals, ethics, and interest in films were highlighted. The research findings indicated that despite a general awareness among educators about imported ideologies, they demonstrated passivity and unwillingness in effectively conveying such awareness to the learners within educational institutions. Babaii and Sheikhi (2017) follow a thorough examination of the textbook episodes and compiled topics. The findings indicated that a substantial majority, surpassing 50%, of the identified subjects exhibit a direct correlation to neoliberal principles, such as employment, financial concerns, interviewing skills, and consumerism. Additionally, an appreciable portion of the topics, categorised as indirect associations, encompasses themes related to celebrities, fashion, and advertising. Xiong and Yuan (2018) utilise a critical discourse analysis approach to investigate the dominant ELT series in China, specifically

the "Go for it!" series, in order to reveal the presence of neoliberal discourse. The results revealed indications of neoliberal principles, such as the promotion of marketisation and individualism, within the locally-produced educational materials.

Bori's (2020) investigation of global English language textbooks aimed to comprehend the representation of neoliberal governmentality within the ELT materials. The study revealed that these textbooks not only replicated neoliberal doctrine but also instilled in learners the values of responsible consumption and entrepreneurship, which are demanded by neoliberal governmentality. The textbooks presented a vision of reality that was permeated by the advantages of entrepreneurship, self-responsibility, corporate social responsibility, and the pleasures of consumption. Jalalian Daghigh and Abdul Rahim (2021) follow the infiltration of neoliberal ideology into domestically produced and globally imported textbooks in Malaysia. They discovered that both collections of textbooks depicted a broad range of neoliberal ideologies. According to the research conducted by Farsani and Rahimi (2022), young learners in Iranian private English language institutes are not exempt from the influence of neoliberalism. The researchers examined a series of textbooks used in these institutes and discovered that they conveyed a hidden curriculum that promoted neoliberal principles such as the importance of 'fame & celebrity,' 'material wealth,' 'consumerism,' and 'competitiveness' to the young Iranian learners.

Park and Lo (2012) state that the concept of neoliberalism has resulted in a shift in perspective regarding language and communication, wherein they are no longer perceived as inherently connected to one's identity. Instead, there is a growing tendency to view them as separable entities that can be manipulated and commodified for economic purposes. The rising perception of English as a crucial ability for future job prospects results in an escalation in the demand for English language services. Consequently, the capacity to afford private English education emerges as a decisive element influencing the life paths of young individuals (Choi & Park, 2013).

Researchers such as Babaii and Sheikhi (2017), Bori and Canale (2022), and Farsani and Rahimi (2023) advocate for the use of critical pedagogy as a means to prevent the dangers posed by neoliberalism. In this regard, Akbari (2008) holds a different opinion from publishers who recommend that coursebook authors should only choose non-controversial topics such as travel, food, and shopping in order to avoid creating content that may lead to disagreements. Freire (1986) believes that critical thinking plays a crucial role in critical pedagogy, which aspires to more than just creating awareness about injustice and questioning societal norms. Its goal is to encourage active engagement in constructing a more equitable society.

Dedicated proponents of critical pedagogy within the realm of education espouse the notion of actively identifying instances of social inequalities and biases, guiding students towards achieving an informed acknowledgement of these inequities, and further empowering them to actively engage in challenging and addressing systems of oppression.

Moreover, these educators exhibit a willingness to adapt their own beliefs and perspectives in order to better align with the needs and viewpoints of their students, as emphasised by Giroux and McLaren (1986) as well as Shor and Freire (1987). According to Ellis (2009) and other researchers, the critical pedagogy approach to English language teaching provides learners with extensive opportunities for learning by creating an environment that is rich in language acquisition. They argue that when learners initiate questions during critical pedagogical tasks, a vital role is played in generating learning opportunities and developing them as active and critical thinkers (Waring, 2009).

Kellner (2000) defines critical pedagogy as analysing education's role in developing individuals and promoting democracy for a just and equal society. It promotes students' critical awareness of structural inequality, political efficacy, and actions to alleviate injustice (Aldana & Byrd, 2015; Seider & Graves, 2020). In recent years, there has been a notable increase in the acknowledgement of the applicability and influence of critical pedagogy within the domain of English language instruction. Within this particular context, critical pedagogy serves as a valuable framework that facilitates the cultivation of learners' critical thinking abilities and their active involvement in exploring topics pertaining to power dynamics, identity construction, and cultural aspects within English-speaking societies. Although there are a number of studies that examine the effect of critical pedagogy in the context of English language education, the existing body of research specifically pertaining to the influence of critical pedagogy on neoliberalism within this context remains insufficient, thereby calling for further inquiry by other scholars in order to address this research gap. As posited by Norton and Toohey (2004), this pedagogical approach underscores the significant role of equipping learners with the capacity for critical analysis and thinking, with the ultimate aim of fostering principles of social justice and equality. During an interview dedicated to the topic of critical thinking, Stephan Brookfield expresses the following statement:

An educator's responsibility is to the student, not to the employing agency, and an educator's responsibility is to understand the internal dynamics of learning, and having that be the logic that drives your actions. Rather than the logic of institutional need. And we're often caught in between the two (Brookfield et al., 2019, p. 83).

## Conclusion

To Harvey (2005), neoliberalism can be characterised as the "financialisation of everything" (p. 33). Savage (2017) claims that the rise of neoliberalism has brought about a fundamental metamorphosis in the operational dynamics of educational institutions. ELT has not remained untouched by this global trend. This paper explores the infiltration of neoliberalism into education systems, with a specific focus on its impact on English Language Teaching. The study highlights how neoliberal ideologies have led to the commodification of education, the transfer of educational responsibility to households, and the emphasis on standardised testing and accountability. The findings demonstrate that neoliberalism

has significantly transformed educational philosophy by prioritising individualism, market-driven principles, and the perception of education as an economic investment.

The study reveals that neoliberalism has had a pervasive influence on education systems globally. It provides examples from Iran, Australia, Europe, and South Korea, showcasing how government schooling systems, curricula, and educational policies have been shaped by neoliberal principles. It also discusses the influence of neoliberalism on ELT, where English is often conceptualised as a marketable skill and individuals are viewed primarily as economic assets. The impact of neoliberalism is evident in ELT textbooks and materials, which often promote themes of employability and consumerism.

These findings have implications for future problems and future studies in education. The study emphasises the need for further academic exploration and critical inquiry to understand the pervasive effects of neoliberalism on education systems and learners. It calls for more research on the consequences of neoliberal governance in education, including its impact on socio-economic inequalities, curriculum design, and the experiences of educators and students. Additionally, future studies could investigate alternative educational paradigms that challenge the neoliberal framework and promote equitable and holistic approaches to education. This article ends with a recommendation by other researchers such as Babaii and Sheikhi (2017), Bori and Canale (2022), and Farsani and Rahimi (2023) that call for a critical pedagogy to avoid threats by neoliberalism.

## References

Abdollahzadeh, E., & Baniasad, S. (2010). Ideologies in the imported English textbooks: EFL learners and teachers' awareness and attitude. *Journal of English Language Teaching and Learning*, 53(217), 1–17.

Ahonen, S. (2002). From an industrial to a post-industrial society: Changing conceptions of equality in education. *Educational Review*, 54, 173–181.

Akbari, R. (2008). Transforming lives: Introducing critical pedagogy into ELT classrooms. *ELT Journal*, 62(3), 30–39. <https://doi.org/10.1093/elt/ccn025>

Aldana, A., & Byrd, C. M. (2015). School ethnic-racial socialization: Learning about race and ethnicity among African American students. *Urban Review*, 47, 563–576. <https://doi.org/10.1007/s11256-014-0319-0>

Andrew, M. (2023). Neo-neoliberalist capitalism, intensification by stealth and campus real estate in the modern university in Aotearoa/New Zealand. *Journal of Applied Learning and Teaching*, 6(2), 393–401. <https://doi.org/10.37074/jalt.2023.6.2.16>

Babaii, E., & Sheikhi, M. H. (2017). Traces of neoliberalism in English teaching materials: A critical discourse analysis. *Critical Discourse Studies*, 15(3), 247–264. <https://doi.org/10.1080/17405904.2017.1398671>

Bacon, C. K., & Kim, S. Y. (2018). "English is my only weapon": neoliberal language ideologies and youth metadiscourse in South Korea. *Linguistics and Education*, 48, 10–21. <https://doi.org/10.1016/j.linged.2018.09.002>

Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228.

Barnawi, O. Z. (2017). *Neoliberalism and English language education policies in the Arabian Gulf*. Routledge.

Bernstein, K. A., Hellmich, E. A., Katznelson, N., Shin, J., & Vinall, K. (2015). Critical perspectives on neoliberalism in second foreign language education. *L2 Journal*, 7(3), 3–14.

Block, D., Gray, J., & Holborow, M. (2012). *Neoliberalism and applied linguistics*. Routledge.

Bori, P. (2020). Neoliberal governmentality in global English textbooks. *Classroom Discourse*, 11(2), 149–163. <https://doi.org/10.1080/19463014.2020.1755327>

Bori, P., & Canale, G. (2022). Neoliberal foreign language education: The search for alternatives. *Critical Inquiry in Language Studies*, 19(4), 307–316. <https://doi.org/10.1080/15427587.2022.2090362>

Bourdieu, R., & Passeron, J. (1977). *Reproduction in education, society and culture*. Sage Publications.

Brookfield, S., Rudolph, J., & Yeo, E. (2019). The power of critical thinking in learning and teaching. An interview with professor Stephen D. Brookfield. *Journal of Applied Learning and Teaching*, 2(2), 76–90. <https://doi.org/10.37074/jalt.2019.2.2.11>

Brown, W. J. (2015). *Undoing the demos: Neoliberalism's stealth revolution*. Zone Books.

Byram, M., & Parmenter, L. (2012). *The common European framework of reference: The globalisation of language education policy*. Multilingual Matters.

Canale, G. (2022). Designing for assessment as recognition of multimodal work in the EAL classroom. In S. Diamantopoulou & S. Orevik (Eds.), *Multimodality in English language learning* (pp. 207–220). Routledge.

Choi, H. J., & Park, J. H. (2013). Historical analysis of the policy on the college entrance system in South Korea. *International Education Studies*, 6(11), 106–121.

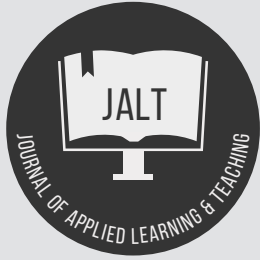
Daghighi, J. A., & Rahim, A. H. (2021). Neoliberalism in ELT textbooks: An analysis of locally developed and imported textbooks used in Malaysia. *Pedagogy, Culture & Society*, 29(3), 493–512. <https://doi.org/10.1080/14681366.2020.1755888>

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.

- Farsani, M. A., & Rahimi, F. (2022). Neoliberalism for young learners in an EFL context: A mixed-methods approach. *Language Culture and Curriculum*, 1–28. <https://doi.org/10.1080/07908318.2022.2110890>
- Freire, P. (1986). *Pedagogy of the oppressed*. Continuum.
- Giroux, H. A. (2008). *Against the terror of neoliberalism: Politics beyond the age of greed*. Paradigm.
- Giroux, H. A., & McLaren, P. (1986). Teacher education and the politics of engagement: The case for democratic schooling. *Harvard Educational Review*, 56, 213–238. <https://doi.org/10.17763/haer.56.3.trr1473235232320>
- Hakala, J., Uusikylä, K., & Jarvinen, E. (2015). Neoliberalism, curriculum development and manifestations of 'creativity.' *Sage Journals*, 18(3), 250-262. <https://doi.org/10.1177/1365480215596239>
- Harvey, D. (2005). *A brief history of neoliberalism*. Oxford University Press.
- Hastings, M. (2019). Neoliberalism and education. *Oxford Research Encyclopedia of Education*. <https://doi.org/10.1093/acrefore/9780190264093.013.404>
- Heller, M. (2003). Globalization, the new economy, and the commodification of language and identity. *Journal of Sociolinguistics*, 7(4), 473–492. <https://doi.org/10.1111/j.1467-9841.2003.00238.x>
- Heller, M. (2010). The commodification of language. *Annual Review of Anthropology*, 39(1), 101–114. <https://doi.org/10.1146/annurev.anthro.012809.104951>
- Heller, M., Duchêne, A. (2012). Pride and profit: Changing discourses of language, Capitaland Nation-State. In A. Duchêne, & M. Heller (Eds.), *Language in late capitalism: Pride and profit* (pp. 1–21). Routledge.
- Hogan, A. (2016). NAPLAN and the role of edu-business: New governance, new privatisations and new partnerships in Australian education policy. *Australian Educational Researcher*, 43(1), 93–110.
- Horiguchi, S., Imoto, Y., & Poole, G. S. (2015). Introduction. In S. Horiguchi, Y. Imoto, & G. S. Poole (Eds.), *Foreign language education in Japan* (pp. 1–18). Sense.
- Husen, T. (1994). *Skola och universitet inför 2000-talet: En utbildningsforskarens perspektiv*. [School and university for the 21st century: An educational researcher's perspective.] Bokförlaget Atlantis.
- Husen, T. (1996). The idea of the university: Changing roles, current crisis and future challenges. In Z. Morsy, & P. G. Altbach (Eds.), *Higher education in an international perspective, critical issues* (pp. 3–20). Garland Publishing.
- Kalalahti, M., & Varjo, J. (2012). Tasa-arvo ja oikeudenmukaisuus perusopetukseen sijoittumisessa ja valikoitumisessa [Equality and fairness in placement and selection in basic education]. *Kasvatus & Aika*, 6(1), 39-55.
- Kellner, D. (2000). Multiple literacies and critical pedagogies. In P. R. Trifonas (Ed.), *Revolutionary pedagogies* (pp. 196–200). Routledge.
- Norton, B., & Toohey, K. (2004). *Critical pedagogies and language learning*. Cambridge University Press.
- Park, J. S. Y. (2016). Language as pure potential. *Journal of Multilingual and Multicultural Development*, 37(5), 453–466. <https://doi.org/10.1080/01434632.2015.1071824>
- Park, J. S. Y., & Lo, A. (2012). Transnational South Korea as a site for a sociolinguistics of globalization: Markets, timescales, neoliberalism. *Journal of Sociolinguistics*, 16, 147–164.
- Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.
- Plehwé, D. (2009). Introduction. In P. Mirowski, & D. Plehwé (Eds.), *The road from Mont Pèlerin* (pp. 1–42). Harvard University Press.
- Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023). A critical perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning and Teaching*, 6(2), 311-331. <https://doi.org/10.37074/jalt.2023.6.2.5>
- Read, J. (2009). A genealogy of homo-economicus: Neoliberalism and the production of subjectivity. *Foucault Studies*, 6, 25–36. <https://doi.org/10.22439/fs.v0i0.2465>
- Savage, G. (2017). Neoliberalism, education and curriculum. In B. Gobby, & R. Walker (Eds.), *Powers of curriculum: Sociological perspectives on education* (pp. 143–165). Oxford University Press.
- Savski, K. (2020). Local problems and a global solution: Examining the recontextualization of CEFR in Thai and Malaysian language policies. *Language Policy*, 19(4), 527–547. <https://doi.org/10.1007/s10993019095398>
- Seider, S., & Graves, D. (2020). *Schooling for critical consciousness*. Harvard Education Press.
- Shin, H. (2016). Language 'skills' and the neoliberal English education industry. *Journal of Multilingual and Multicultural Development*, 37(5), 509–522. <https://doi.org/10.1080/01434632.2015.1071828>
- Shor, I., & Freire, P. (1987). What is the 'dialogical method' of teaching? *Journal of Education*, 169, 11–31. <https://doi.org/10.1177/002205748716900303>
- Smith, M. S. (2022). Social reproduction as language policy: The neoliberal co-option of English in global Japan. *Educational Policy*, 36(7), 1652–1678. <https://doi.org/10.1177/0895904821999840>

- Urciuoli, B. (2008). Skills and selves in the new workplace. *American Ethnologist*, 35(2), 211–228. <https://doi.org/10.1111/j.1548-1425.2008.00031.x>
- Vandrick, S. (2014). The role of social class in English language education. *Journal of Language, Identity & Education*, 13(2), 85–91.
- Waring, H. Z. (2009). Moving out of IRF (initiation-response-feedback): A single case analysis. *Language Learning*, 59(4), 796–824.
- West, G. F. (2019). Navigating morality in neoliberal spaces of English language education. *Linguistics and Education*, 49, 31–40. <https://doi.org/10.1016/j.linged.2018.12.004>
- Xiong, T., & Yuan, Z. M. (2018). “It was because I could speak English that I got the job”: Neoliberal discourse in a Chinese English textbook series. *Journal of Language, Identity & Education*, 17(2), 103–117. <https://doi.org/10.1080/15348458.2017.1407655>

Copyright: © 2024. Ehsan Nikpouya and Gholamreza Zareian. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Metrics in research impact assessment and grant funding: Insights from researchers in the "Reviewer 2 Must Be Stopped!" Facebook group

Valentine J. Owan <sup>A</sup>	A	<i>Department of Educational Foundations, University of Calabar, Calabar, Nigeria; Ultimate Research Network (URN), Calabar, Nigeria</i>
Victor U. Agama <sup>B</sup>	B	<i>Department of Educational Foundations, University of Calabar, Calabar, Nigeria; Ultimate Research Network (URN), Calabar, Nigeria</i>
John O. Odey <sup>C</sup>	C	<i>Department of Science Education, Federal University Wukari, Nigeria</i>
Delight O. Idika <sup>D</sup>	D	<i>Institute of Education, University of Calabar, Calabar, Nigeria</i>

### Keywords

Funding strategies;  
grant allocation;  
interdisciplinary research;  
performance metrics;  
research impact;  
research impact assessment.

### Abstract

Research assessment and grant funding are vital to higher education. However, the reliance on quantitative metrics in these processes has raised concerns about their validity and potential negative consequences. This study aims to investigate the game of numbers in research assessment and grant funding, focusing on the perspectives of experienced researchers from around the globe. Accidental sampling elicited responses from more than 15 experienced researchers across different academic disciplines, institutions, and countries. The data were collected from the popular "Reviewer 2 Must be Stopped!" Facebook platform, which includes more than 135,000 members across the globe. Two posts were made, allowing participants to share their experiences, perspectives, and concerns related to metrics and numbers in research assessment and grant funding. The results from the thematic analysis revealed diverse perspectives among experienced researchers. Some participants expressed concerns about the dominance of quantitative metrics, highlighting the limitations and potential biases associated with their use. Others acknowledged the value of certain indicators in showcasing research impact. Moreover, the impact of metrics on grant funding awards was also documented. The study highlights the necessity for a more balanced and context-aware approach to research assessment and grant funding, incorporating qualitative measures and acknowledging the diverse nature of research impact.

### Correspondence

owanvalentine@gmail.com <sup>A</sup>

### Article Info

Received 4 September 2023  
Received in revised form 30 January 2024  
Accepted 25 February 2024  
Available online 4 March 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.19>

## Introduction

In the ever-evolving landscape of higher education, research has become the nucleus driving institutions toward excellence, innovation, and societal advancement. Pursuing knowledge, discovery, and groundbreaking solutions propels universities and researchers to transcend boundaries, inspiring a relentless quest for meaningful contributions to their respective fields. As academia expands, the significance of research impact assessment and securing grant funding has become an intricate game, intertwining ambition, intellect, and a strategic understanding of the dynamics that govern this ever-competitive arena. Universities and research institutions across the globe strive to advance the frontiers of science, drive societal development, and address the most pressing challenges facing humanity (Owan et al., 2023; Odigwe & Owan, 2022). Central to this pursuit is the ever-increasing emphasis on research impact assessment and securing grant funding, which have become critical barometers of success in today's competitive academic milieu (Lambovska & Todorova, 2021; Owan & Asuquo, 2022). In fact, academics are now considered fictionally, as manager, zombie, ninja, nervous wreck, activist, third space precariat, and early career precariat (Andrew, 2023) due to their academic roles; some of which appear to involve irrational compliance and obedience to modern academic demands.

As the 21st century unfolds, the global scientific community has grappled with an unprecedented explosion of information thanks to advancements in technology and communication. This rapid proliferation of knowledge has led to immense opportunities and profound challenges. While scholars have the potential to access a vast repository of information, the sheer volume of research output makes it increasingly challenging to discern impactful and reliable research from the deluge of mediocre work (Odigwe & Owan, 2022). Consequently, in the past, institutions and funding agencies have used qualitative assessment methods to evaluate the influence of research outcomes and gauge their real-world relevance (Bakker et al., 2020). Qualitative evaluations, including peer reviews and expert judgments, strive to provide a more nuanced understanding of the societal implications and practical applications of research (Louder et al., 2021; Reed et al., 2021).

However, in recent years, quantitative metrics have been used in research impact assessments (Fleming et al., 2021; Lauronen, 2020; Ma & Ladisch, 2019). Journal impact factors, citation counts, h-indices, and altmetrics, among other numeric indicators, have emerged as quantitative indicators for attempting to capture the reach and significance of individual research papers and researchers (Owan & Owan, 2021). See Moed and Halevi (2015) for a detailed discussion of these and other metrics. In fact, as Professor Stephen Brookfield submits, "everything in higher education has been commodified to some degree" (Brookfield et al., 2019, p.84). The commodification of research output does not imply that they would not serve helpful purposes. However, many scholars have questioned the validity of using metrics to assess the quality of scholarly output, stating that the numbers tell nothing but mere fragments of the complex and multifaceted reality of academic research (Calò, 2022; Hicks et

al., 2015; Wilsdon, 2016). They argue that reducing the value of scholarly work to a set of quantitative measures not only oversimplifies the true impact of research but also promotes a culture of academic conformity, where researchers might prioritise publishing in high-impact journals over pursuing groundbreaking and potentially transformative studies (Xu & Li, 2016). It has been documented that heavy reliance on metrics is responsible for several unethical practices, such as salami slicing, 'abeg put my name syndrome', citation cartels, gambling and h-index manipulation, such as 'cite-me-I-cite-you', unnecessary self-citation inflation and other unacceptable practices (Moed & Halevi, 2015; Owan & Asuquo, 2022; Owan & Owan, 2021).

Complicating matters further, substantial differences exist in publication and citation practices across subject fields. For instance, in molecular biology, cited reference lists in scientific publications tend to be longer and more focused on recent articles than do those in fields such as mathematics, resulting in higher citation rates for target articles in the former, especially during the early postpublication period (Moed & Halevi, 2015). This divergence can affect the accuracy of absolute citation counts, making normalised indicators, such as those comparing citation impact to the world citation average in the relevant subfields, more appropriate. However, employing normalised indicators is not without its challenges. Bibliometric research has revealed complexities in counting variations in institutions and individual names, potential errors due to limited database coverage, and the exclusion of "gray literature", such as technical reports, which may lead to partial assessments in certain disciplines (Donner et al., 2020; Guerrero-Bote et al., 2021; Jappe, 2020).

Furthermore, the prestige of the journal where a paper is published introduces another layer of intricacy. Various indices, such as the Impact Factor (Web of Science), Source Normalised Impact per Paper (Scopus), and SCImago Journal Rank (SJR), are utilised to gauge journal prestige, each addressing specific challenges in research assessment (Owan et al., 2023). The debate over the fairness and accuracy of citation indices as measures of productivity and impact persists. Despite such opposition and evident limitations, citation counts remain widely used in research assessments (Moed & Halevi, 2015; Owan & Owan, 2021). While citation analysis can be useful, it should complement other evaluative approaches to provide a more comprehensive and well-rounded assessment of research impact and productivity. The need for an objective measurement and the failure of metrics in "telling the whole story" about a given research work has prompted scholars to seek new ways of assessing research impact. For instance, subjective tools such as questionnaires have been developed to measure research impact (e.g., Dembe et al., 2014; Solans-Domènech et al., 2019).

Moreover, while evaluating "large cancer research funding in Australia", Bowden et al. (2018) focused on key indicators, such as knowledge production, career advancement, generation of newer tools for future research, further income generation, development of newer policies and products, and other health, social and behavioural benefits. Similarly, during an evaluation, Ravenscroft et al. (2017) found that the results of metrics used as measures of

research impact did not conform well with the results of the Research Excellence Framework (REF). The authors suggested that the non-academic impact of research be evaluated using information mined from a broad range of resources, including social media engagement, news articles and political debates arising from academic work. Clements et al. (2017) introduced “snowball metrics,” a more robust methodology that promises to improve upon the current system, but concluded that data from quantitative assessments of research impact are informative but should never replace human judgments in peer reviews when assessing research quality.

Measuring research impact depends on how the concept is defined and contextualised. Research impact could have different meanings in academic and broader socioeconomic contexts. There is a disparity between the academic and broader socioeconomic impacts of research and assessments in the UK considering these two dimensions separately (Penfield et al., 2014). The term ‘impact’ refers to the fact that the influence of research has advanced beyond academia (Chowdhury et al., 2016). Nevertheless, this distinction is unclear in impact assessments outside the UK, where academic outputs and socioeconomic impacts are often viewed as one. In Nigeria, research impact is considered holistically, and individuals are assessed based on its utility value in academia, with little emphasis on socioeconomic impact or innovativeness. Some universities in Nigeria now regard mere publishing in journals indexed in Scopus and Web of Science as indicators of research impact due to the international reputation of the two databases (Owan et al., 2023).

Since metrics are now a part of the research assessment system, it is crucial to understand whether they matter in deciding who receives grant funding and their overall role in the impact assessment of research. Research assessment and grant funding are important components of higher education, as they enable universities and institutions to support research undertakings and academic advancement (Sato et al., 2021). The increased emphasis on metrics and quantitative indicators has created a complex and competitive landscape that governs research assessment and grant funding in higher education (Hicks, 2012). This has led to various assessment frameworks with unique indicators and criteria (Reed et al., 2021). The “game of numbers” has increasingly influenced decision-making, with researchers, institutions, and funding bodies using metrics extensively to evaluate academic performance, allocate resources, and make funding decisions.

According to the findings of Thuna and King (2017), respondents expressed concerns about how metrics impacted their appointment to editorial boards, selection for administration, grant funding, evaluation of other scholars for promotion, job applications, and choices of publication venues. After reviewing a large pool of previously funded projects, Győrffy et al. (2020) found moderate positive correlations between the scientometric standing (such as the h-index, citation counts and yearly average) of principal investigators during grant submission and their future research output.

The existing body of research on the role of metrics in grant funding decisions and research impact assessment is limited and lacks sufficient empirical attention. Despite extensive discussions and commentary on the topic (e.g., Adam et al., 2018; Carpenter et al., 2014; Helmer et al., 2020; Moed & Halevi, 2015; Recio-Saucedo et al., 2022; Wilsdon, 2016), studies investigating this topic are scarce. This knowledge gap in the literature suggests the need for comprehensive and rigorous research to better understand how metrics influence funding decisions and impact assessments in the academic and research community. Several systematic reviews (e.g., Cruz-Rivera et al., 2017; Milat et al., 2015; Peter et al., 2017; Razmgir et al., 2021) have described a lack of qualitative studies in this area, creating a methodological gap. These studies have recommended that future studies on this subject adopt a qualitative approach to assess the role of metrics in grant funding decisions and research impact assessment. Qualitative studies can provide in-depth insights into the perceptions, experiences, and attitudes of scholars, funding agencies, and policymakers, which may not be fully captured by quantitative analyses alone (Yadav, 2022). It is important to analyse participants’ views to enhance research quality, inform policy guidelines, and minimise unintended consequences. The insights gained from this study can lead to improved funding allocation, better recognition of interdisciplinary research, and greater compliance with metrics. For these reasons, the present study was conceived to draw qualitative insights into the role of research metrics in grant funding and impact assessment. Specifically, the study assessed the following:

1. The role of metrics in research funding decisions and practices across different disciplines and regions;
2. The role of metrics in research impact assessment in higher education institutions.

### **Research questions**

The following research questions were answered in this study:

1. How has metric utilisation influenced research funding decisions and practices across different disciplines and regions?
2. How have metrics affected research impact assessment in higher education institutions?

### **Methods**

#### **Research design**

The research philosophy guiding this study was interpretivism. It acknowledges that reality is socially constructed and seeks to understand the diverse perspectives and experiences of experienced researchers regarding research assessment and grant funding processes. A qualitative research approach was employed for this study. This approach allowed the researchers to delve into the subjective experiences and perceptions of the participants, providing a comprehensive understanding of their views on metrics in research



assessment and grant funding. The chosen research strategy was a qualitative case study. This strategy facilitated an in-depth exploration of the phenomenon of interest (the game of numbers in research assessment and grant funding) within the context of the "Reviewer 2 Must be Stopped!" Facebook group. This study was conducted as cross-sectional research, collecting data from the Facebook group over two occasions. The data collection process spanned two months from June to July 2023 to ensure sufficient data diversity and representation. The research design of this study is illustrated using Saunder's research onion (Saunders et al., 2009, 2012) in Figure 1.

pool of experienced researchers with relevant perspectives. Two separate posts were made in the group to facilitate data collection. The first post invited participants to share their experiences regarding how reliance on metrics has influenced their access to grant funding at their respective institutions. The second post sought to understand how metrics are used in research impact assessment across different academic institutions. Participants' responses were collected and transcribed verbatim for analysis. The data collection process was meticulously documented to promote replicability.

### Verifiability and trustworthiness

To enhance the verifiability and trustworthiness of the study, triangulation was employed. The researchers cross-referenced the data from multiple sources, such as posts and comments, to validate the findings and ensure consistency. The research team engaged in intercoder reliability checks, wherein two independent researchers coded a subset of the data to establish coding agreement. Any discrepancies were resolved through discussion, ensuring the reliability of the coding process. Member checking was also conducted, wherein a select group of participants was asked to review and validate the preliminary findings. Incorporating their feedback strengthened the study's validity and ensured an accurate representation of the participants' perspectives.

### Ethical consideration

The research adhered to strict ethical guidelines throughout the study. Informed consent was obtained from all participants, who provided detailed information about the study's purpose, data usage, and rights. Participant anonymity was maintained using pseudonyms, and all identifying information were removed from the transcribed data. The study obtained approval from the Institutional Review Board (IRB) to ensure compliance with ethical standards. Ethical documentation, including informed consent forms and IRB approval, were meticulously maintained for transparency and replicability.

### Data analysis

Thematic analysis was used to analyse the collected data. The researchers immersed themselves in the transcribed responses to identify recurring themes and patterns. A systematic coding process was used to categorise and organise the identified codes into overarching themes, ensuring a comprehensive understanding of the participants' experiences and perspectives. The data analysis process was thoroughly documented, including detailed explanations of coding decisions and theme development. This documentation served to promote replicability and transparency in the study's findings.

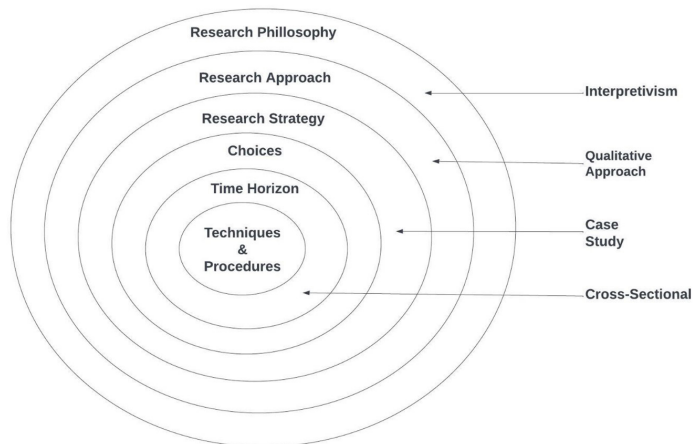


Figure 1: Research onion of Saunder showing the research design of this study.

### Participants

Accidental sampling was utilised to select more than 15 experienced researchers from various academic disciplines, institutions, and countries. Participants were chosen based on their active engagement in the Facebook group's research assessment and grant funding discussions. All participants were considered experts with substantial knowledge of their respective institutions' research assessment and grant funding processes. Table 1 describes the demographic profiles of the participants.

Table 1: Demographic profiles of the participants.

Participant	Label	Gender	Discipline	Location
Participant 1	P1	Male	Chemistry	Scotland
Participant 2	P2	Male	Postdoctoral Fellow	Finland
Participant 3	P3	Female	Human Ecology	Sweden
Participant 4	P4	Female	Researcher	Lithuania
Participant 5	P5	Male	Human Geography	Italy
Participant 6	P6	Male	Biology	India
Participant 7	P7	Male	Researcher	Europe
Participant 8	P8	Male	Physiology	Pakistan
Participant 9	P9	Female	Neuroscience	Mexico
Participant 10	P10	Male	Computational Biology	USA
Participant 11	P11	Male	Management	USA
Participant 12	P12	Male	Data Science	France
Participant 13	P13	Male	Biology Education	Hungary
Participant 14	P14	Female	Educational Measurement	Nigeria
Participant 15	P15	Male	Internal Medicine	Nigeria

### Data collection

The data were collected from the "Reviewer 2 Must be Stopped!" Facebook Group, which boasts more than 135,000 academic members worldwide. The group's popularity and strict regulation of academic membership ensured a rich

## Results

The results of the qualitative analysis demonstrate a diverse range of perspectives regarding the role of metrics in modern day research-related practices. The results are organised according to the major themes that emerged from the analysis:

### Theme 1: Metrics in funding evaluation

Participants' views on the significance of metrics in research funding decisions were explored in this theme. On one hand, some participants acknowledge that their institutions and funding agencies rely heavily on metrics when making funding decisions. Metrics such as publication counts, impact factors, and citation indices are viewed by some participants as objective measures of researchers' productivity and impact. Some participants disclosed that their institutions rely on metrics when making research funding decisions. However, some expressed concerns about potential flaws in the system, such as oversimplification and the disregard of qualitative contributions. Moreover, some participants shared the common view that even though metrics are extensively used for research funding, there is growing awareness of the need to be more cautious about using this approach for evaluation. Some participants revealed the following regarding the role of metrics in funding evaluation:

P1: "I am not sure about securing funding, but my university management is obsessed with them in terms of 'performance management'."

P7: "From my experience (and my experience only), it is very important in Europe."

P9: "Depends on the country and area. In my country (Mexico), in theory, funding should not depend on metrics but rather on quality. In practice, committees base their decisions mainly on metrics, even when they were specifically instructed to ignore them. It is an old custom that is hard to let go."

P10: "I think this goes by the economic concept that one thing is as good as its alternatives allow it to be. There are always more PIs (Principal Investigators) than funding, and when it comes to rationing time, how else do we objectively compare them other than things like impact factors and publications? By assigning a board of 'morally incorruptible true experts' and centralised planning Soviet style? I do not know a better solution than the current system."

On the other hand, several participants argue against the overreliance on metrics in grant funding. They highlight the importance of valuing the strength of research proposals and other qualitative aspects over publication metrics. Participants mention a shift toward prioritising well-written and thoughtful research proposals rather than excessively emphasising metrics. This shift suggests a growing recognition of the limitations of metrics and the need for a more balanced evaluation process that considers the potential impact and novelty of research projects. On the

opposing end, some participants revealed the following regarding the role of metrics in grant funding:

P2: "I am sure there are huge differences between fields, countries, and funding agencies here (I am in Biogeosciences in Finland). My primary funding agency, the Research Council of Finland, has banned impact factors from CVs/publication lists because the ranking of a journal is a very poor predictor of the impact of an individual article - even many Nature papers turn out to have very few citations. We include our individual citation record (e.g., total number of citations, h-factor). My overall impression is that funding agencies have a more qualitative and holistic view of research performance, i.e., they look at your CV. Publication record, supervision record, project management experience, previous funding, public engagement on the research topic, etc., are all considered."

P3: "Metrics have not influenced my funding or applications. In my small transdisciplinary field, they would not make much sense."

P4: "The use of metrics for assessment depends on the funder and the people who will review your application. I think ERC now focus on the quality of work and proposal rather than indices."

P6: "Simple. Those who do not understand research by reading it have no other way but rely on indices. Those who know science will not. However, such people are not found in funding agencies nowadays."

P8: "In our part of the globe, Pakistan, no weightage is given to these metrics such as citations, h-index, etc., while considering Research Grant funding. The grant reviewers mostly pay attention towards novelty and practical applicability of the project."

P5: "Since the use of metrics for evaluation lacks scientific and analytical validity, only those who continue to advocate for their use can provide answers to this question. Is the practice of incorporating Impact Factors (IFs) of the journals where people published being implemented? That would be irrational, though it is possible (I have come across CVs with cumulative IFs, which is simply absurd). Are they relying on IFs as indicators of the quality of individual articles instead of reading them? It may be a careless approach, but it is feasible. In summary, there is no reasonable way to utilise metrics effectively, and this inquiry delves into imprudent and flawed practices."

In a follow-up discussion, Participant 5 was further asked to share a perspective on circumstances in which every decision about hiring, firing or promotion was an invitation to accuse bias and whatever best suits the candidate; having a set of numbers from disinterested sources to back the decision can be very useful. The response obtained is as follows:

P5: "Apologies for being frank, but believing that "numbers" are free from biases is precisely that insanity I was referring to before. Are we talking of kids playing with things they do not understand or of competent adults? Can an adult with PhD and academic credentials do something they know is sloppy and not be blamed? Let us be adults. Having metrics that appear to be objective and independent of personal bias allows them to make decisions that put the best people into jobs, grant resources where they will be used most productively, or get rid of non-performing personnel without themselves being targeted."

## **Theme 2: Critique of metric-based evaluation**

Theme 2 focused on participants' critiques and reservations regarding the use of metrics in research funding assessments and decision-making. The participants expressed their concerns and scepticism about the reliance on metrics as the primary or sole criterion for evaluating research proposals and allocating funding. These studies lend their attention to metric-based evaluation methods' limitations and potential drawbacks. Participants expressed a growing concern about the limitations and potential adverse effects of relying heavily on quantitative measures to allocate research funding. The perspectives of five participants exemplify the reservations surrounding the use of metrics in funding decisions.

P1: "My university management are obsessed with them in terms of 'performance management'. However, there is uncertainty about whether metrics directly impact funding decisions in my institution."

P7: "I know mediocre scientists who have accumulated a good number of articles and citations by being at the right time in the right lab, and they can secure funding for projects which are either extremely incremental or downright stupid. I have the counterexample of an Indian postdoc who is brilliant but does not have an impressive CV and struggles to secure even personal grants. The reason scientific progress is slowing down so much is, in my opinion, a combination of a lack of investments coupled with a system that fosters mediocrity to some extent. I am part of this system and have not done much to change it. Things are so bad, but if we voice it, we are discredited since we have no power. Moreover, those with power will not change things because why would they?"

P9: "In my institution, there is a contradiction between the stated emphasis on quality over metrics in funding decisions and the reality that committees still heavily rely on metrics. This discrepancy ushers in challenge in breaking away from an entrenched culture of metric-based evaluation, despite acknowledging the importance of focusing on research quality."

P10: "Metrics are used in my institution due to the sheer number of Principal Investigators competing for limited funding. Funding agencies are now facing

a dilemma in efficiently comparing researchers objectively. This approach is not appropriate, especially now that many agencies are adopting inclusive approaches to research impact assessment."

P15: "Before the advent of metrics, assessment of scholarly impact was primarily based on qualitative evaluations, peer reviews, and expert judgments rather than quantitative measures. The impact was also based on the number of patents earned through discoveries that are new, non-obvious, industrially applicable, disclosed in detail, and fall within the scope of patent-eligible subject matter. However, with metrics, such as citation counts and journal impact factors, the assessment of scholarly impact shifted towards more quantitative and standardised approaches. These do not have any meaning for me, but my institution continues to emphasise that people go for Q1- and Q2-indexed journals in Scopus and Web of Science. I do not understand why universities like ours will emphasise these things over our quality of work. Even teaching is less important during assessment than metrics (ordinary numbers). As a result, many people continue to fail promotion even though they have given their best for their universities through teaching, research and community service. We should go back to the old assessment system, or instead of relying solely on metrics, researchers should be made to justify their research's social and economic value as part of the promotion requirements."

The critique presented by these participants underscores several common concerns about using metrics in research funding assessments. These include the risk of oversimplifying research productivity, fostering a focus on quantity rather than quality, perpetuating a culture of mediocrity, and creating an environment where researchers are incentivised to "game the system" to improve their metrics artificially. Moreover, the critiques point to a potential disconnect between the espoused values of funding decisions based on quality and the actual reliance on metrics in practice. This discrepancy suggests the need for a more balanced approach to research funding evaluation that considers both quantitative metrics and qualitative evaluations of research impact and potential.

## **Theme 3: Metrics and research impact assessment**

This theme focuses on participants' views regarding the role of metrics during research impact assessment. Participants were asked to share their views on how metrics are used during their promotion career advancement appraisals and measure their impact in their respective fields. One of the key findings is the varied perspectives on the significance of metrics in research impact assessment. While some participants acknowledged the importance of metrics, others criticised their use as scientifically unsound and difficult to justify in practice. Nevertheless, some participants agreed that metrics are widely used for research impact assessment at their institutions.

P12: "In my country, research impact assessment is based on lists provided by the government. This is very problematic, given how political ideology can skew evaluations."

P13: "Hungary in general: papers in Scopus indexed journals (preferred: Q1, Q2), number of citations, h-index. In all disciplines, however, specifics (e.g., how many citations or papers are needed for a promotion) vary according to discipline. In humanities/social sciences, monographs also count (but they do not replace journal papers)."

P14: "My institution relies so much on metrics such as the h-index of scholars and impact factors of the journals where authors published their papers. Authors must meet the minimum criteria set in the metric-based system to be promoted to different ranks. For instance, you will only be promoted to a professor only if you have at least five papers in Scopus or Web of Science, other publications in national or association journals, and a h-index of 5 or higher. Those seeking promotion to the rank of associate professor needed approximately 3 papers in these databases and a h-index of 3. The total number of publications must be sufficient to give you the required points. Although other aspects, such as the number of courses taught, students' research supervised, and conferences attended, are necessary conditions that must be satisfied, research metrics seem to have a domineering place among all these measures. For instance, you do not even get a clearance to submit your records for promotion assessment if you do not meet the Scopus or h-index requirements. Besides, papers published in journals with impact factors are graded more favourably with higher scores than those without them."

On the other hand, some respondents agreed that while metrics are being used, there is a gradual shift toward complementing metric-based assessments with subjective evaluations. This can be seen in the direct quotations of two participants:

P8: "The Impact Factor of a journal and h-index of a scholar plays a pivotal role in research assessment and hence his/her promotions in academic/research institutes. However, the tide is now turning against these contentious metrics, and emphasis is gradually being levied upon the Impact of Research, its applicability, and its practicality."

P11: "At my institution, we use multiple measures of research impact, including number of citations, quality of journal, journal impact factor, external assessments of your scholarship, research being used in doctoral seminars, awards, select membership due to your areas of expertise, and invitations extended to you to give presentations/talks based on your research. I enjoy it because there are different avenues for different scholarships to show impact."

## Discussion

The first research question sought to explore the perspectives of scholars on the role of metrics in research funding decisions. The qualitative analysis revealed a diverse range of perspectives among the participants. Some participants acknowledged the extensive use of metrics, such as publication counts, impact factors, and citation indices, as objective measures of researchers' productivity and impact. Some participants expressed concerns about how funding decisions are based on numeric indicators when deciding who gets what. However, concerns about potential flaws in the system, such as oversimplification and neglect of qualitative contributions, were also voiced. This finding corroborates the evidence presented by Dinsmore et al. (2014) that funding agencies such as the 'Wellcome Trust' utilise metrics in making funding decisions and consider altmetrics an innovative approach. These findings also align with the results obtained by Thuna and King (2017), where participants stated, among other things, that impact metrics were important for them to secure grants. These findings further support the findings of Györfy et al. (2020) that the scientometric standing of an author was important for grant funding and future research output. Nevertheless, the approaches adopted in the cited studies and the present study were quite different.

Many participants acknowledged the importance of research metrics but suggested that a more balanced approach to evaluation be used, where metric-based evidence is supplemented with subjective evaluation inputs from core domain experts. This finding implies that policymakers should consider incorporating qualitative methods alongside metrics to improve the validity and reliability of research funding decision-making. This shift indicates a move toward a more comprehensive evaluation approach that values qualitative aspects, such as research impact and novelty. This aligns with the results of Butler et al. (2017), who found that no perfect all-encompassing metric exists for measuring research impact and that no single traditional metric can accommodate all facets of research impact in the modern era. By prioritising the strength of research proposals and qualitative contributions, higher education institutions and funding agencies can encourage researchers to focus on producing high-quality and innovative research beyond mere publication counts. This approach may foster a research culture that rewards excellence and creativity, leading to more impactful and groundbreaking research outcomes. In addition, the findings of this study have implications for teaching and learning in higher education institutions, where the focus is gradually shifting from the classroom activities of academics towards research. Using a balance approach will ensure that academics are assessed by cumulative scores from both their classroom and research performance, not just from the latter, as we currently see today.

The critique of overreliance on quantitative metrics, which may foster a culture of mediocrity and artificial inflation of metrics, calls for caution when using metrics as the primary evaluation criterion. A balanced evaluation approach considering quantitative metrics and qualitative impact is crucial to mitigate these potential drawbacks. Encouraging

researchers to produce high-quality work with genuine societal relevance can promote a research ecosystem that values excellence and meaningful contributions. These findings corroborate the results of Bakker et al. (2020), who revealed that participants familiar with metrics expressed concerns about their misuse and desire to be involved in decision-making around their use. Similarly, in an editorial published by ACS Nano, a group of experts resolved that research impact measurement should go beyond citations and publications, considering real-world effects and encouraging thoughtful assessment by funding agencies, institutions, and researchers while fostering incentives for research development and contributions beyond academia (Chai et al., 2022).

The second research question explores participants' views on the role of metrics in research assessment in higher education institutions. Participants expressed diverse views on the importance of metrics in research evaluation. While some acknowledged the pivotal role of metrics such as the Impact Factor and h-index in research assessment and promotions, others criticised their use as scientifically unsound and challenging to justify. This finding aligns with the results of Deeming et al. (2018), where participants provided broad support for using standardised and customised metrics in research impact assessment. This diversity of perspectives underscores the complexity of using metrics as evaluation criteria in research impact assessment. The implications of these findings are twofold: first, the diverse views on the importance of metrics underscore the need for a balanced approach to research evaluation that considers both quantitative metrics and qualitative assessments; second, the criticisms of metric-based evaluation emphasise the necessity of addressing the limitations and potential biases associated with metrics to ensure a fair and comprehensive assessment of research impact in higher education institutions.

This shift involves using multiple measures of research impact, including citations, journal quality, external assessments, awards, and research usage in doctoral seminars. This finding agrees with the results of Deeming et al. (2018), where most participants felt that the current research environment encourages academics to focus on publishing papers and building their academic reputation, which sometimes clashes with making broader impacts outside of academia. Emphasising qualitative aspects alongside metrics promotes a fairer and more accurate assessment, encouraging researchers to focus on producing meaningful and impactful research outcomes. This finding implies that a one-size-fits-all evaluation system may not accurately capture diverse contributions and research priorities across different regions and disciplines. By considering these contextual factors, policymakers and institutions can mitigate potential biases and design evaluation frameworks that are fair, transparent, and inclusive. This approach ensures that researchers' achievements are assessed and aligned with their respective regions' and disciplines' specific characteristics and goals, promoting a more equitable and effective research evaluation process.

## **Limitations and prospective research directions**

While this study provided valuable insights into the influence of metrics on research funding decisions and assessments, it also had several limitations that should be acknowledged. First, the sample size was small and not fully representative of all disciplines and regions. This may limit the generalisability of the findings and might not fully capture the diverse perspectives on metric utilisation in different academic contexts. Another limitation is the potential for self-selection bias among the participants. Those who chose to participate in the study may have had stronger opinions about the topic, leading to a skewed representation of views.

Additionally, the qualitative nature of the study might limit the ability to quantify and measure certain aspects of the participants' responses, making it challenging to draw precise conclusions. Furthermore, the study focused on participants' perspectives and experiences, which might not entirely reflect the actual practices and policies of funding agencies and institutions. Incorporating data from funding agencies and institutional records could provide a more comprehensive picture of how metrics are used in decision-making.

In future research, it would be beneficial to conduct larger-scale studies with more diverse samples to increase the generalisability of the findings. Longitudinal studies tracking changes in metric utilisation over time help identify trends and patterns. Moreover, combining qualitative and quantitative methods could offer a more robust analysis of the impact of metrics on research funding decisions. Finally, exploring the perspectives of key stakeholders, such as funding agencies, university administrators, and policymakers, would provide a more comprehensive understanding of how metrics are integrated into funding and assessment practices. Despite these limitations, the study lays the groundwork for further exploration and calls attention to the need for balanced and nuanced evaluation methods in research funding decisions. By addressing these limitations and pursuing future research, we can continue to advance the knowledge in this critical area and promote more effective and equitable research evaluation practices in higher education institutions.

## **Implications for teaching and learning**

The implications of this study for teaching and learning in higher education systems are far-reaching, especially in redefining the balance among the three statutory duties of academics—teaching, research, and community service. The study suggested the need for academic institutions to de-emphasise metrics when assessing scholarly contributions. This shift can recalibrate the priorities of scholars, prompting them to strike a more balanced approach between teaching and research. As pressure to meet strict quantitative metrics diminishes, academics can allocate more time and energy to excel in their teaching responsibilities. The current emphasis on research metrics may have led to a sense of insensitivity towards teaching. With a reduced reliance on metrics, scholars can redirect their focus toward improving the quality of teaching. This can lead to the development of innovative teaching methods, a stronger emphasis on

student engagement, and a dedication to fostering a positive learning environment. This study implies that the pursuit of metrics sometimes results in the production of papers lacking practical relevance. Shifting away from a solely metric-driven approach allows scholars to prioritise research with genuine societal impact. This, in turn, can lead to more practical and applicable research outcomes, enriching the learning experience for students by connecting theoretical knowledge with real-world scenarios. The third statutory duty, community service, often takes backseat in the pursuit of research metrics. As scholars pay more attention to teaching and research with societal relevance, community services can be better integrated into academic agendas. This may involve active engagement with local communities, sharing expertise with the public, and contributing to solutions for real-world problems. De-emphasising metrics in assessment allows space for a more holistic academic experience. Students can benefit from educators who are not solely focused on research output but are equally invested in creating a positive learning environment and making meaningful contributions to the community. The intense pressure to meet research metrics can contribute to burnout among academics. By shifting the focus away from metrics and fostering a more balanced approach, institutions can contribute to the well-being of scholars. This, in turn, positively impacts the teaching and learning environment, as educators can bring a more positive and energised mindset to their work.

## Conclusion

This study significantly contributes to the field by highlighting the diverse perspectives among scholars regarding the role of metrics in research funding decisions and assessments in higher education institutions. The findings revealed a diverse range of perspectives among participants, with some acknowledging the significance of metrics as objective measures of researchers' productivity and impact. In contrast, others criticised the overreliance on metrics and advocated for a more comprehensive evaluation approach. Moreover, the shift toward adopting multiple measures of research impact in certain institutions reflects a growing recognition of alternative indicators beyond traditional metrics. The findings underscore the necessity of a balanced evaluation approach that integrates qualitative assessments with traditional quantitative metrics. This approach is critical for funding agencies to reconsider their evaluation criteria. The recommendation to adopt a balanced evaluation approach, which values qualitative contributions and societal impact alongside traditional metrics, holds the potential to reshape the research landscape. This shift can foster a culture that rewards excellence and creativity, encouraging researchers to produce high-quality and innovative work beyond mere publication counts. The implications of this research underscore the importance of a balanced and thoughtful approach to research evaluation that considers both quantitative metrics and qualitative aspects of research quality. As institutions and funding agencies strive to make informed decisions about research funding and assessment, these insights can guide the development of more effective and fair evaluation methods that capture the true impact and novelty of research contributions. Additionally, the

study emphasises the need for ongoing dialogue and reflection on using metrics in research assessment to foster a culture of rigorous and unbiased evaluation in the academic community. Based on the findings and conclusions of this study, the following recommendations were made:

1. Institutions should work toward creating evaluation frameworks that go beyond quantitative metrics. Qualitative assessments, such as peer reviews and expert evaluations, should be incorporated to provide a more holistic view of scholarly contributions, including teaching and community service.
2. Higher education institutions should encourage scholars to develop a balanced academic portfolio that encompasses teaching, research, and community service. They should recognise and reward contributions in all three areas, fostering a culture that values well-rounded academics.
3. Tertiary institutions should offer professional development opportunities that focus on enhancing teaching skills and promoting impactful community engagement. This could include workshops, training sessions, and mentorship programs to support academics in all aspects of their roles.
4. Institutions should revisit promotion and tenure criteria to ensure that they reflect a balanced assessment of faculty contributions. We should consider weighting teaching and community services alongside research on promotion decisions to encourage a more equitable distribution of effort.
5. Higher education institutions should actively seek and consider student feedback on the effectiveness of teaching. This can provide valuable information concerning the impact that educators have on students' learning experiences and help in refining teaching methods.
6. Higher education institutions should educate faculty, administrators, and stakeholders about the limitations and potential biases associated with relying solely on quantitative metrics.
7. Funding agencies should consider incorporating a balanced evaluation approach that goes beyond quantitative metrics. While metrics such as publication counts, impact factors, and citation indices provide valuable insights, funding decisions should also consider qualitative contributions and the societal impact of research.
8. Funding agencies should encourage researchers to provide evidence of their engagement with communities, the practical relevance of their

work, and contributions to teaching during promotion assessment. This approach ensures that funding decisions consider the broader spectrum of scholarly activities, fostering a research culture that values excellence in various dimensions.

## References

- Adam, P., Ovseiko, P. V., Grant, J., Graham, K. E. A., Boukhris, O. F., Dowd, A.-M., Balling, G. V., Christensen, R. N., Pollitt, A., Taylor, M., Sued, O., Hinrichs-Krapels, S., Solans-Domènech, M., & Chorzempa, H. (2018). ISRIA statement: Ten-point guidelines for an effective process of research impact assessment. *Health Research Policy and Systems, 16*(1), Article No. 8. <https://doi.org/10.1186/s12961-018-0281-5>
- Andrew, M. B. (2023). Come to the cabaret: Voices from the modern university. *Journal of Applied Learning & Teaching, 6*(2), 17–27. <https://doi.org/10.37074/jalt.2023.6.2.19>
- Bakker, C., Cooper, K., Langham-Putrow, A., & McBurney, J. (2020). Qualitative analysis of faculty opinions on and perceptions of research impact metrics. *College & Research Libraries, 81*(6), 896–912. <https://doi.org/10.5860/crl.81.6.896>
- Bowden, J. A., Sargent, N., Wesselingh, S., Size, L., Donovan, C., & Miller, C. L. (2018). Measuring research impact: A large cancer research funding programme in Australia. *Health Research Policy and Systems, 16*(1), 39. <https://doi.org/10.1186/s12961-018-0311-3>
- Brookfield, S. D., Rudolph, J., & Yeo, E. (2019). The power of critical thinking in learning and teaching. An interview with Professor Stephen D. Brookfield. *Journal of Applied Learning and Teaching, 2*(2), 76–90. <https://doi.org/10.37074/jalt.2019.2.2.11>
- Butler, J. S., Sebastian, A. S., Kaye, I. D., Wagner, S. C., Morrissey, P. B., Schroeder, G. D., Kepler, C. K., & Vaccaro, A. R. (2017). Understanding traditional research impact metrics. *Clinical Spine Surgery, 30*(4), 164–166. <https://doi.org/10.1097/BSD.0000000000000530>
- Calò, L. N. (2022). Impact metrics and science evaluation. *Revista Peruana de Medicina Experimental y Salud Pública (Peruvian Journal of Experimental Medicine and Public Health), 39*, 236–240. <https://doi.org/10.17843/rpmpesp.2022.392.11171>
- Carpenter, C. R., Cone, D. C., & Sarli, C. C. (2014). Using publication metrics to highlight academic productivity and research impact. *Academic Emergency Medicine, 21*(10), 1160–1172. <https://doi.org/10.1111/2Facem.12482>
- Chai, Y., Chen, X., Fan, H. J., Lau, S. P., Li, S., Liu, B., Wong, W.-Y., Zhao, N., & Zheng, Z. (2022). Beyond metrics. *ACS Nano, 16*(8), 11485–11486. <https://doi.org/10.1021/acsnano.2c07662>
- Chowdhury, G., Koya, K., & Philipson, P. (2016). Measuring the impact of research: Lessons from the UK's research excellence framework 2014. *PloS One, 11*(6), Article No. e0156978. <https://doi.org/10.1371/journal.pone.0156978>
- Clements, A., Darroch, P. I., & Green, J. (2017). Snowball metrics – Providing a robust methodology to inform research strategy – But do they help? *Procedia Computer Science, 106*, 11–18. <https://doi.org/10.1016/j.procs.2017.03.003>
- Cruz-Rivera, S., Kyte, D. G., Aiyegbusi, O. L., Keeley, T. J., & Calvert, M. J. (2017). Assessing the impact of healthcare research: A systematic review of methodological frameworks. *PLoS Medicine, 14*(8), Article No. e1002370. <https://doi.org/10.1371/journal.pmed.1002370>
- Deeming, S., Reeves, P., Ramanathan, S., Attia, J., Nilsson, M., & Searles, A. (2018). Measuring research impact in medical research institutes: A qualitative study of the attitudes and opinions of Australian medical research institutes towards research impact assessment frameworks. *Health Research Policy and Systems, 16*(1), 1–20. <https://doi.org/10.1186/s12961-018-0300-6>
- Dembe, A. E., Lynch, M. S., Gugiu, P. C., & Jackson, R. D. (2014). The translational research impact scale: Development, construct validity, and reliability testing. *Evaluation & the Health Professions, 37*(1), 50–70. <https://doi.org/10.1177/0163278713506112>
- Dinsmore, A., Allen, L., & Dolby, K. (2014). Alternative perspectives on impact: The potential of ALMs and altmetrics to inform funders about research impact. *PLoS Biology, 12*(11), Article No. e1002003. <https://doi.org/10.1371/journal.pbio.1002003>
- Donner, P., Rimmert, C., & van Eck, N. J. (2020). Comparing institutional-level bibliometric research performance indicator values based on different affiliation disambiguation systems. *Quantitative Science Studies, 1*(1), 150–170. [https://doi.org/10.1162/qss\\_a\\_00013](https://doi.org/10.1162/qss_a_00013)
- Fleming, P., Rudolph, J., & Tan, S. (2021). 'Never let a good crisis go to waste'. An interview with Professor Peter Fleming on dark academia, the pandemic and neoliberalism. *Journal of Applied Learning and Teaching, 4*(2), 110–120. <https://doi.org/10.37074/jalt.2021.4.2.14>
- Guerrero-Bote, V. P., Chinchilla-Rodríguez, Z., Mendoza, A., & de Moya-Anegón, F. (2021). Comparative analysis of the bibliographic data sources dimensions and Scopus: An approach at the country and institutional levels. *Frontiers in Research Metrics and Analytics, 5*, Article No. 593494. <https://doi.org/10.3389/frma.2020.593494>
- Györfy, B., Herman, P., & Szabó, I. (2020). Research funding: Past performance is a stronger predictor of future scientific output than reviewer scores. *Journal of Informetrics, 14*(3), 101050. <https://doi.org/10.1016/j.joi.2020.101050>
- Helmer, S., Blumenthal, D. B., & Paschen, K. (2020). What is meaningful research, and how should we measure it? *Scientometrics, 125*(1), 153–169. <https://doi.org/10.1007/s11192-020-03649-5>

- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251-261. <https://doi.org/10.1016/j.respol.2011.09.007>
- Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. <https://doi.org/10.1038/520429a>
- Jappe, A. (2020). Professional standards in bibliometric research evaluation? A meta-evaluation of European assessment practice 2005–2019. *PLoS One*, 15(4), Article No. e0231735. <https://doi.org/10.1371/journal.pone.0231735>
- Lambovska, M., & Todorova, D. (2021). 'Publish and flourish' instead of 'Publish or perish': A motivation model for top-quality publications. *Journal of Language and Education*, 7(1), 141–155. <https://doi.org/10.17323/jle.2021.11522>
- Lauronen, J. P. (2020). The dilemmas and uncertainties in assessing the societal impact of research. *Science and Public Policy*, 47(2), 207-218. <https://doi.org/10.1093/scipol/scz059>
- Louder, E., Wyborn, C., Cvitanovic, C., & Bednarek, A. T. (2021). A synthesis of the frameworks available to guide evaluations of research impact at the interface of environmental science, policy and practice. *Environmental Science & Policy*, 116, 258-265. <https://doi.org/10.1016/j.envsci.2020.12.006>
- Ma, L., & Ladisch, M. (2019). Evaluation complacency or evaluation inertia? A study of evaluative metrics and research practices in Irish universities. *Research Evaluation*, 28(3), 209-217. <https://doi.org/10.1093/reseval/rvz008>
- Milat, A. J., Bauman, A. E., & Redman, S. (2015). A narrative review of research impact assessment models and methods. *Health Research Policy and Systems*, 13, 1-7. <https://doi.org/10.1186/s12961-015-0003-1>
- Moed, H. F., & Halevi, G. (2015). Multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology*, 66(10), 1988-2002. <https://doi.org/10.1002/asi.23314>
- Odigwe, F. N., & Owan, V. J. (2022). University collaborative research and wealth creation. In J. A. Undie, J. B. Babalola, B. A. Bello, & I. N. Nwankwo (Eds.), *Management of higher educational systems* (pp. 995–1002). University of Calabar Press. <https://bit.ly/3Oi630I>
- Owan, V. J., & Asuquo, M. E. (2022). "Publish or perish," "Publish and perish": The Nigerian experience. In J. A. Undie, J. B. Babalola, B. A. Bello, & I. N. Nwankwo (Eds.), *Management of higher educational systems* (pp. 986–994). University of Calabar Press. <https://bit.ly/3TEHEDu>
- Owan, V. J., Basse, B. A., & Ubi, I. O. (2023). Construction and standardisation of an instrument measuring lecturers' persistence to publish in Scopus-indexed journals. *Journal of Applied Learning & Teaching*, 6(2), 158–171. <https://doi.org/10.37074/jalt.2023.6.2.37>
- Owan, V. J., & Owan, M. V. (2021). Complications connected to using the impact factor of journals for the assessment of researchers in higher education. *Mediterranean Journal of Social & Behavioral Research*, 5(1), 13–21. <https://doi.org/10.30935/mjosbr/10805>
- Penfield, T., Baker, M. J., Scoble, R., & Wykes, M. C. (2014). Assessment, evaluations, and definitions of research impact: A review. *Research Evaluation*, 23(1), 21-32. <https://doi.org/10.1093/reseval/rvt021>
- Peter, N., Kothari, A., & Masood, S. (2017). Identifying and understanding research impact: A review for occupational scientists. *Journal of Occupational Science*, 24(3), 377-392. <https://doi.org/10.1080/14427591.2016.1277547>
- Ravenscroft, J., Liakata, M., Clare, A., & Duma, D. (2017). Measuring scientific impact beyond academia: An assessment of existing impact metrics and proposed improvements. *PLoS One*, 12(3), Article No. e0173152. <https://doi.org/10.1371/journal.pone.0173152>
- Razmgir, M., Panahi, S., Ghalichi, L., Mousavi, S. A. J., & Sedghi, S. (2021). Exploring research impact models: A systematic scoping review. *Research Evaluation*, 30(4), 443-457. <https://doi.org/10.1093/reseval/rvab009>
- Recio-Saucedo, A., Crane, K., Meadmore, K., Fackrell, K., Church, H., Fraser, S., & Blatch-Jones, A. (2022). What works for peer review and decision-making in research funding: A realist synthesis. *Research Integrity and Peer Review*, 7(1), 1-28. <https://doi.org/10.1186/s41073-022-00120-2>
- Reed, M. S., Ferré, M., Martin-Ortega, J., Blanche, R., Lawford-Rolfe, R., Dallimer, M., & Holden, J. (2021). Evaluating impact from research: A methodological framework. *Research Policy*, 50(4), Article No. 104147. <https://doi.org/10.1016/j.respol.2020.104147>
- Sato, S., Gyax, P. M., Randall, J., & Schmid Mast, M. (2021). The leaky pipeline in research grant peer review and funding decisions: Challenges and future directions. *Higher Education*, 82(1), 145-162. <https://doi.org/10.1007/s10734-020-00626-y>
- Saunders, M. N. K., Lewis, P., & Thornhill, A. (2009). Understanding research philosophy and approaches to theory development. In M. N. K. Saunders, P. Lewis, & A. Thornhill (Eds.), *Research methods for business students* (pp. 122-161). Harlow, England: Pearson Education. <https://bit.ly/3NZGj9u>
- Saunders, M. N. K., Lewis, P., & Thornhill, A. (2012). *Research methods for business students* (6th ed.). Harlow, England: Prentice Hall. <https://bit.ly/43tqaiw>
- Solans-Domènech, M., Pons, M. V. J., Adam, P., Grau, J., & Aymerich, M. (2019). Development and validation of a questionnaire to measure research impact. *Research Evaluation*, 28(3), 253-262. <https://doi.org/10.1093/reseval/rvz007>
- Thuna, M., & King, P. (2017). Research impact metrics: A faculty perspective. *Partnership: The Canadian Journal of*



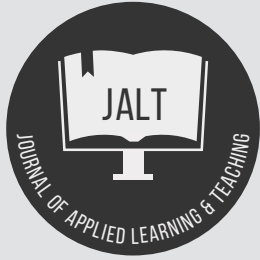
*Library and Information Practice and Research*, 12(1), 1–25. <https://doi.org/10.21083/partnership.v12i1.3906>

Wilsdon, J. (2016). *The metric tide: Independent review of the role of metrics in research assessment and management*. Sage Publications. <https://doi.org/10.4135/9781473978782>

Xu, F., & Li, X. (2016). The changing role of metrics in research institute evaluations undertaken by the Chinese Academy of Sciences (CAS). *Palgrave Communications*, 2(1), 1-6. <https://doi.org/10.1057/palcomms.2016.78>

Yadav, D. (2022). Criteria for good qualitative research: A comprehensive review. *The Asia-Pacific Education Researcher*, 31(6), 679-689. <https://doi.org/10.1007/s40299-021-00619-0>

Copyright: © 2024. Valentine J. Owan, Victor U. Agama, John O. Odey and Delight O. Idika. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## A content analysis of tweets on toxic doctoral supervision

Oluchi Ojinamma  
Okere<sup>A</sup>

A

*Albert Ilemobade Library, Federal University of Technology Akure, Nigeria*

---

### Keywords

Content analysis;  
doctoral students;  
PhD supervision;  
toxic PhD supervisors;  
Twitter.

---

### Abstract

Doctoral students are expected to conduct independent research and produce original contributions to their field of study. Therefore, doctorate programmes are rigorous and demanding, and they require a significant amount of dedication and hard work. High attrition and dropout rates generally mark the post-coursework phase of the programme because of the difficulties that go with the independent research aspect of the study. Supervisory practices are identified as major reasons for the discontinuance of these programmes. The purpose of this study was to identify the nature of communication on toxic supervision of PhD students by Twitter users. The data collection consisted of tweets posted between January 1, 2020, and March 12, 2023. The methodology used was content analysis, which included the examination of the tweets for themes and trends collected within the time frame. It provides valuable data on the lived experiences of PhD students under toxic academic supervision. This study reflects the value of Twitter as a tool for research and as a medium of expression and emotional support for PhD researchers. The study will contribute to policy-making and training in supervisory practices.

---

### Correspondence

ojinammaatu@gmail.com <sup>A</sup>

---

### Article Info

Received 9 September 2023  
Received in revised form 7 January 2024  
Accepted 7 March 2024  
Available online 19 March 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.26>

## Introduction

The purpose of a PhD programme is to provide students with the opportunity to undertake original research in a specific field of study and make a significant contribution to the existing body of knowledge. PhD programmes are designed to equip students with advanced research skills, critical thinking, and analytical abilities to become independent researchers and scholars in their field. During the dissertation phase, a supervisor is assigned to the student to offer guidance in creating an original research output in the form of a thesis or dissertation. Thus, "a professional contract akin to an apprenticeship" is established between the two (Jabre et al., 2021).

Doctoral programmes span between three and four years, and are usually completed in two main phases, the coursework component of about one academic year, and the dissertation component of about two years. It is typical for students to complete the coursework part of their postgraduate studies on time, but then struggle for years to finish their thesis (Costa, 2018). This difficulty leads some students to give up altogether, resulting in a lack of formal qualification after many years of wasted rigour and stress. Across universities, there is a high rate of PhD students' dropout, sometimes between 36% and 51% (Young et al., 2019; Payne, 2021). Effective supervision is crucial to the success and quality of postgraduate education. Several studies have shown that the quality of supervision can have a significant impact on timely completion rates, research quality, student satisfaction, and retention rates (Alam et al., 2013; Gruzdev et al., 2020).

The findings of this study hold significant implications for policy-making and training initiatives aimed at improving supervisory practices within academia, ultimately contributing to the overall well-being and success of doctoral students.

## Literature review

The working relationship between the supervisor and supervisees is well-researched due to its strategic role in the success rate of PhD programmes. Supervisors hold explicit authority which is typical of hierarchical power dynamics. Power imbalances inherent in these relationships can lead to complications, thus, a nuanced understanding of the power dynamics between PhD students and their supervisors is desirable. Brookfield et al. (2022) highlight Foucault's (2000) crucial insights into power dynamics within academia and their relevance for ethical teaching. Foucault's analysis reveals the pervasive nature of power, particularly through disciplinary and bio-power. These could turn lifelong learning into a daunting experience akin to a 'lifelong nightmare'.

Effective supervision of students helps ensure that they are developing the necessary knowledge, skills, and attitudes to meet the requirements of their research project and achieve their final goals. A toxic or abusive working relationship between supervisors and their supervisees would have a direct impact on students' capacity to complete the many tasks that culminate in the completion of the programme and

graduation. Educational institutions often resemble tightly controlled systems, mirroring prisons in their constraining structures. Thus, teachers must be vigilant of power dynamics among both themselves and students, as differences in status and privilege are imported into the classroom environment. Foucault's insights support the imperative for educators to navigate power dynamics conscientiously to uphold ethical teaching practices (Brookfield et al., 2022).

In examining the lived experiences of doctoral students, Al Makhmreh and Stockley (2020) review the nature of mentorship and identify three levels, namely: authentic mentorship, average mentorship, and toxic mentorship. Students who experienced inadequate or detrimental mentorship were likely to feel stressed and emotionally drained.

A toxic relationship is characterised by high levels of physical or emotional stress, lack of communication, bullying, a lack of respect for boundaries, and a lack of appreciation in the working relationship. In a toxic relationship, the unhealthy nature of the power dynamics is generally observable. In a thematic analysis of the behavioural characteristics of toxic research supervisors, Shahnawaz and Siddiqi (2022) taxonomised the traits as oppressive management style, misuse of authority, inadequate guidance, and erratic emotions (primary characteristics); and inadequate capacity to interact, low level of competence and high level of prejudice and bias (secondary characteristics).

Similarly, Gruzdev et al. (2020) carried out a survey involving PhD students at top-rated Russian universities to categorise supervision styles and examine their impact. Using cluster analysis, six styles were identified: superheroes, hands-off supervisors, research advisors, dialogue partners, research practice mediators, and mentors. Hands-off supervisors provide little guidance, resulting in the lowest student satisfaction and longest expected time-to-degree. However, many students with hands-off supervisors were still satisfied, suggesting a "disengagement compact" between students and supervisors. The styles with the highest satisfaction and shortest expected time-to-degree were superheroes and mentors, who provide managerial and expert support. The students indicated that performing administrative and advisory functions was critical for effective supervision and student progress.

Submissive individuals in toxic relationships often experience feelings of being unappreciated, misinterpreted, undervalued, and sometimes even subjected to mistreatment. Such relationships can harm the mental and emotional well-being of the subordinate individual. Al Makhmreh and Stockley (2020) suggest that doctoral students can complete their studies despite facing toxic mentorship, but this achievement may come at the cost of their mental and emotional health. Shahnawaz and Siddiqi (2022) identify low levels of self-disclosure and a poor sense of identification with their peers as some of the characteristics displayed by doctoral students in India.

Aside from the supervisor-student relationship, supervisory practices may be affected by other factors. For instance, supervisors may have difficulties with the intellectual and

psychological components of the postgraduate programme. They may lack sufficient research knowledge and abilities to facilitate supervision. There may also be a mismatch between the student and his supervisor (Priyadarshini et al., 2022; Muraraneza et al., 2020).

As further noted by Priyadarshini et al. (2022), supervisors are often overburdened by academic, administrative, and organisational responsibilities, leaving little time for effective engagement during supervision meetings. Supervisors' lack of time, absence from the institution, poor supervisory techniques, and a high supervisor-student ratio may impact student and supervisor productivity. Also, students' low internal motivation, poor time management skills, or weak academic writing skills are other challenges associated with students' non-performance, and an increase in student attrition in PhD programmes.

To add to these situations, the nature of PhD programmes is changing globally. Massification and internalisation are some of the emerging challenges in universities that affect doctoral scholarship (Gruzdev et al., 2020). In recent times, universities are increasingly facing challenges related to diversity, inclusion, and retention of students from diverse backgrounds. Also, universities have started embracing new paradigms, such as online supervision. These have an impact on the supervisory role of faculty members in postgraduate programmes. According to Bogelund (2015), currently, the market-driven approach dominates, possibly affecting the quality of research and supervisor job satisfaction. These changes are engendered by cultural changes in higher education.

Neoliberalism has also upset academic culture. The pervasive influence of neoliberalism is contributing to a toxic environment characterised by individualisation, competition, and the commercialisation of knowledge. This toxic culture has become the new norm, leading to heightened pressures and frustrations among academics, managers, and students alike (Moore et al., 2021; Andrew, 2023). Consequently, the university landscape is fraught with unsustainable work relationships and a proliferation of toxic behaviours. In the face of these challenges, scholars are confronted with the erosion of academic freedom, collegiality, and traditional university culture.

Tepper et al. (2017) identified three drivers of supervisory abuse of subordinates in work settings. The factors that are not mutually exclusive include social learning, identity threat, and self-regulatory impairment. Supervisors may perceive that their attitudes towards their subordinates are socially acceptable and rewarding. They may also have strong individual identities, which make them demonstrate their superiority over others and display high personal sensitivity to threats. Finally, supervisors' self-regulatory impairments that may promote abusive behaviours include work stress, poor sleep quality and exercise, and surface acting.

Hazell et al. (2020) conducted a meta-analysis to examine the mental health of PhD researchers and found that PhD students have a higher prevalence of mental health difficulties compared to the general population, which is a global phenomenon. The study identified several individual, interpersonal, and systemic factors that contribute to

mental health problems among PhD students. Among these, isolation and identification as female were the most significant risk factors, along with being single, not having children, and having a lower economic status. The study also found that common psychiatric disorders, such as depression and anxiety, were prevalent among PhD students. Additionally, the quality of the supervisory relationship was found to be essential for maintaining a positive workplace environment.

Masek and Alias (2020) describe fit as the most essential requirement for effective supervision. The fits described are fit in expectation, fit in thinking, and fit in personality and style. The implication is that both actors must have similar mindsets to work together effectively, as a good fit is the key to the excellent interpersonal working relationship between the research student and his advisor.

Muthanna and Alduais (2021) investigated the relationship between research supervision and research integrity, laying the onus of promoting ethical behaviour and integrity in supervision squarely on the shoulders of the supervisor. The supervisor is expected to provide justice, fidelity, autonomy, beneficence, and non-maleficence when relating to supervisees; also, assuming the role of a research supervisor entails taking ethical responsibility for the conduct and output of their students' research.

Twitter content provides a rich source of data that can be analysed for insights into public opinion and sentiments. It is becoming more than a mere tool for marketing or advertising. Ferreira (2021) notes that Twitter could function as a research tool that can support the postgraduate training process. Twitter provides access for both students and supervisors to have access to discipline-specific and interdisciplinary discussions, advice, and collaborations on a global level. The usefulness of Twitter spans various aspects of digital doctoral tradition, including enculturation, communities of practice, and research identity for both students and supervisors. It has also positioned itself as a tool for scholarly (peer) exchange. It also has significant value for research applications (Chen et al., 2021). As a result, it is an excellent source of information for identifying the latent powerplay in the academic system (Liu & Woo, 2021).

The goal of this study is to analyse Twitter content on the subject of toxic supervisor-supervisee relationship to understand its nature, and its effect on student performance satisfaction and progress of the PhD programmes. Investigating toxic supervision in doctoral programmes addresses a critical issue that significantly impacts the success and well-being of PhD students. The study is positioned within the research field of doctoral education and supervision. It contributes to the discourse on effective supervisory practices and holds implications for policy-making for higher education, and training initiatives aimed at improving the overall quality of postgraduate education.

## Method and data

### Methodology

The study method adopted is content analysis. A content analysis is a quantitative approach to qualitative data obtained from communication mediums. It is a method of systematic evaluation of documents or oral communication that enables the researcher to make “replicable and valid inferences by interpreting and coding textual materials.” Through content analysis, unstructured data can be simplified, and trends, patterns, and intentions of the contents of communication in audio, video, pictorial and textual formats can be analysed and interpreted for better understanding.

### Data collection and extraction

Tweets related to toxic PhD supervision relationships from January 1, 2020, to March 12, 2023, were extracted, representing the most recent tweets, three years prior to the research and post-COVID-19. Three keywords, “PhD supervisor”, “toxic PhD supervisor”, and “abusive PhD supervisor” were used to mine the content of Twitter using Twitter’s API (Application Programming Interface). The advanced search function of Twitter was also used to widen the search to include 2020 data. The tweet objects (such as tweet text, publishing date, media, and URLs) were extracted using the tweet IDs of the Twitter-API. The total dataset consisted of 368 related tweets and replies. After extracting tweets with the search terms, the data was scanned to identify and remove tweets that were not contextually relevant to the topic of interest. This resulted in a total of 172 tweets selected for qualitative analysis.

The inclusion and exclusion criteria included: (i) relevant tweets may include a visual (image or video) and (ii) relevant tweets must be in English (both image and tweet text). Two research assistants coded and compared the results. Thereafter, some revisions were made to the coding to refine it based on their findings.

### Coding and data analysis

Coding was done manually. The Speech Act Theory, developed by philosophers like J. L. Austin and expanded by John Searle categorises linguistic expressions into speech acts based on the speaker’s intentions and their impact on the listener. Searle identifies five main categories: declarations, assertives, expressives, directives, and commissives (Searle, 1979). These speech acts play a vital role in shaping communication beyond literal meanings (Barrero, 2023). Understanding these categories helps analysts interpret the intentions behind utterances, including those observed in tweets on platforms like Twitter. This model was used to provide clarity on the nature of the posts. A WordIt word cloud was used to provide a summarised visual representation of the text data. The tool was used to highlight the most prominent words within the datasets with the aim of aiding the understanding of the underlying themes and trends present in the data.

## Ethics

Since the data being collected is publicly available, the researcher did not seek informed consent from individuals. However, I ensured that I did not collect any sensitive information, such as personal information or direct mentions that could harm individuals’ or institutions’ reputations or violate their privacy.

### Research questions

The research questions that guided the study are:

- i. How is toxicity expressed in the tweets?
- ii. What are the prevalent experiences of toxic supervision identified in the tweets?
- iii. What are the consequences of toxic supervision?
- iv. What are the prevalent keywords?
- v. What major themes emanate from the content?

## Results

Table 1: Nature of the tweets.

Nature of Tweet	Speech Act	Frequency	Percentage
Complaints	Expressives	58	33.70%
Suggestions/Advice	Directives	43	25%
Opinion	Assertives	37	21.5%
Questions	Directives	12	7%
Humour	Expressives	10	5.8%
Encouragement	Expressives	9	5.2%
Others		3	1.8%
<b>Total</b>		<b>172</b>	<b>100%</b>

Table 1 explains how toxicity is expressed in the tweets. A frequency count of the tweets revealed that most of the tweets were complaints about the experiences of the PhD students or mentions of the negative experiences of other PhD students (33.7%). Similarly, suggestions or pieces of advice were given by a high percentage of tweet posters (25%). About 22% of the respondents made comments that could be considered opinions on issues concerning toxic PhD supervision. Based on Searle’s Speech Act Classification (1979), the tweets were distributed into three speech acts: expressives (45%), directives (32%), and assertives (22%). There is an obvious non-representation of commissive and declaration speech types. The dominance of expressives suggests that Twitter serves as a medium for individuals to vent their grievances and seek emotional support from their peers. The presence of directive and assertive speech acts, such as suggestions/advice and opinions, highlights that individuals are also using Twitter to offer guidance, share opinions, and provide insights into dealing with toxic PhD supervision. This indicates a level of engagement and activism within the community to address and mitigate issues related to toxic supervision. The absence of commissive and



Table 4 shows the themes that emanate from the content, while Figure 2 shows the themes presented as a chart in percentages.

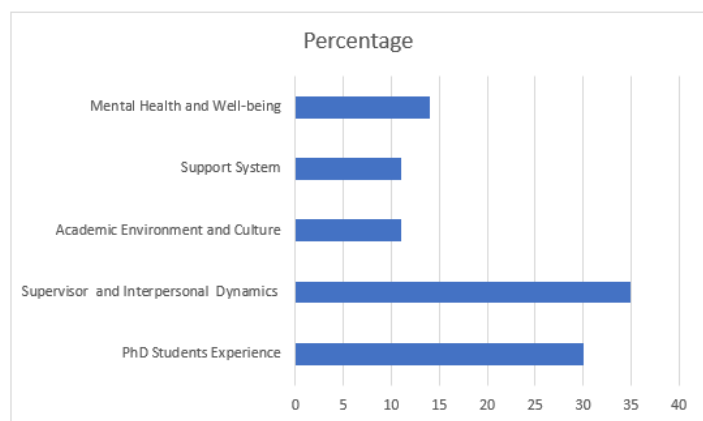


Figure 2: Themes expressed in the Tweets.

The graph above illustrates the themes expressed in the Tweets. The discourse on toxic PhD supervision revolves around five major themes: PhD student experiences, supervisor selection and interpersonal dynamics, academic environment and culture, mentorship and support systems for PhD students, and mental health and well-being. These themes were derived through a coding and recoding process. Most of the tweets were themed around supervisor and interpersonal dynamics (35%) and PhD students lived (negative) experiences (30%). These five themes form a comprehensive corpus of aspects that should be considered during research, planning and/or interventions.

## Discussion

From the findings, PhD students consulted with "Academic Twitter" on the topic mostly for "venting" or pouring out their emotions (expressive act), and to provide answers to those seeking support (directive). Tweets like ... *During my MPhil (supposed to be a PhD, My supervisor called us idiot and stupid almost every week. He also likes to use toxic words to blame us such a XXX, low IQ, etc. This made us to have a severely low self-esteem, zero confidence and being afraid to give an argument* are good examples of this expressive category. The predominance of expressive speech acts reflects the emotional toll and frustration experienced by individuals facing toxic PhD supervision. Accordingly, a commensurate number of tweets reflected suggestions or advice for example, *Future doctoral students your supervisor more important than university prestige or Don't fall for the "prestige" trap*, opinions and encouragement. The absence of commissive and declarative acts may suggest a hesitancy or lack of concrete action towards addressing systemic issues within academia. This suggests little confidence and the need for further dialogue, advocacy, and collective action to address and reform practices related to doctoral supervision. Malik et al. (2019) and Liu and Woo (2021) confirm that Academic Twitter serves the important role of community management. Suggestions included advising, reporting, and quitting. There were tweets that specifically called out institutions where they perceived these toxic

cultures prevailed. The tweets reflect the multifarious roles of Twitter, especially for emotional support.

Further, the study shows that toxic PhD manifested as the narcissistic behaviours of some supervisors, including intentionally wasting students' time, underrating students, and exploiting them by asking them to undertake tasks outside their academic requirements. The tweets reflect experiences *like my supervisor scoffed and said that sleep deprivation was part of PhD and ... I was humiliated by a senior professor... in front of all my peers ....* O'Hara and Cook (2018) report on these types of microaggressions meted out to students, including that they engage in activities beyond their academic expectations by their supervisors, are assumptions and insensitivity about social class background, invalidation of cultural experiences and identities, pressure to assimilate the dominant cultural norms, insensitive remarks about financial circumstances and institutional barriers and policies. The high workload forced on some PhD students and demands reflected in the tweets, *I don't take all of my annual leave, and therefore, neither should you and also, ... He even asked me to do his conference presentation slides etc. I'm being tortured mentally.* Some students were mandated to spend between 65 - 80 hours per week in their labs. The high attrition rate and long completion period could also be blamed on the waste of time doing "nothing", while the supervisor ignored students. Gorup and Laufer (2020) reflect many of these narcissistic and oppressive tendencies of supervisors.

Other themes emerged as well. The study showed that institutions and other colleagues are complicit in creating or sustaining some of these unprofessional and unethical cultures in a number of ways: they look on, even when they are aware of the experiences students are undergoing; they do not provide effective support systems; and in fact, protect the culpable staff members: *...He's known for problems with students, but brings in \$ so nothing is done...;...even if I report misconduct, they're neglected or punished; ...Unfortunately calling them out doesn't work out in the student's favour....*

Sexism, misogyny, and racial biases are part of the unpalatable experiences of some postgraduate students under supervision. The effect of biases and other prejudices on the well-being and academic progress of students is profound. These types of biases are very dangerous to the academy:

*One thing that I learned from my experience as the only female student of an advisor is that often, no matter what you do and how well you do it, you will not be as worthy of your advisors' time. It is called misogyny. We should not have to dance around trying to work around it....*

Three dimensions highlighted by Brookfield et al. (2002) as ethical, productive, and responsible exercise of power by an educator include first, authentic facilitation of student learning, even if it means redirecting or challenging their approaches. Secondly, transparent communication of the rationale behind exercising power, with constant disclosure of the reasons behind decisions and actions, and finally, ongoing opportunities for students to critique the exercise of power, allowing for feedback and reflection to address

any perceived issues of arbitrariness or unfairness directly.

In their study, Gorup and Laufer (2020) and Kis et al. (2022) discuss the effect of poor supervisory practice on the prospects of doctoral students. The findings of the study show that many of the students with toxic supervisors tended to quit academics entirely, change labs or institutions, and develop low levels of confidence and mental health problems as expressed in the tweets: *Sometimes quitting is the only way forward; ... which unfortunately became toxic causing me to change supervisor at a critical time (ended up not finishing my PhD...; ... I get a lot of imposter syndrome and insecurity following a really bad experience; I developed anxiety and depression because of how I was treated during my PhD and have spent thousands of dollars on therapy.*

In conclusion, the themes that emerged from the study include interpersonal dynamics, PhD students' experiences, mental health and well-being, academic environment and culture and support systems. In the model developed by van Rooij et al. (2021), a key predictor of PhD candidates quitting the programme was the quality of their relationship with their supervisor. Specifically, a lower-quality relationship was associated with a higher likelihood of contemplating leaving. The study also identified other influential factors, including project-related aspects such as autonomy, workload, and alignment with the supervisor's research, which emerged as significant new predictors in the final model.

Similarly, the findings of the study about the role of interpersonal dynamics as a significant aspect of the supervisor-supervisee relationship align with Dericks et al.'s (2019) report. Supervisory supportiveness, rather than academic qualities such as research record or reputation, emerged as the primary predictor of PhD student satisfaction. Departmental academic qualities and supportiveness were also significant determinants, giving credence to the value of academic culture and environment. In contrast, peer group factors appeared less influential. These determinants exhibited consistency across different disciplines and countries, according to this international study.

Supervisory practices can be improved through training. Haven et al. (2022) showed that a 3-day training involving responsible research practices (RRPs) and interpersonal skills, resulted in improved supervision skills, as reported by both the PhD students and the supervisors. Also, based on Chugh et al.'s (2022) model, problems associated with supervisory feedback which involve the content, processes and expectations of the feedback must be tackled holistically and synergistically with respect to the three actors: institutions, supervisors and students.

## Conclusion

Toxic supervision practices have been identified as a significant reason for students discontinuing their doctoral programmes. However, doctoral students' experiences under toxic academic supervision remain insufficiently interrogated. This study sheds light on this issue, based on the analysis of Twitter data, which exposes a lot of discrete but negative information concerning the toxic supervision

of PhD students.

The findings of this study show that PhD students generally experience a high level of stress, which emanates from interpersonal interactions with their supervisors and some systemic factors. These are also expressed as narcissistic, exploitative behaviours and attitudes from the supervisor. These stressors lead to responses such as premature quitting of programmes, mental health challenges, and delays in the completion of their programme. The study confirms the true nature of the power dynamics in the academy. It also confirms that Twitter is a robust platform for emotional support for PhD students, given its potential to provide a community for individuals to be inspired, encouraged, and advised on many aspects of their lived experiences.

## Limitations of the study and suggestions for future research

A major limitation of this study is that the data comes solely from Twitter, which may not be representative of the overall population of PhD students. This approach may not offer a comprehensive representation of the entire population of PhD students, as it only captures the experiences of those who choose to share their experiences on Twitter. Also, since cultural differences play a role in interpersonal relationships, it would be more responsive to explore the topic of PhD supervision on a spatial basis.

Future research endeavours could adopt a mixed-methods approach that integrates Twitter (social media) data with interviews or surveys of PhD students to expand upon the current findings and address these limitations. A combination of quantitative and qualitative methods to assess the prevalence of various supervision-related issues would provide a balanced result on the subject.

Furthermore, comparing the perceptions of students, advisors, and administrators could offer a more holistic perspective on the strategies required to address toxic supervision. To gain a deeper understanding of the developments and changes in supervision experiences, longitudinal data and location-based data collection could be employed to investigate how policies and interventions could influence reported supervision encounters over time and in various regions.

## Recommendations

1. Social media spaces as safe spaces: Students and their supervisors should explore Academic Twitter to garner knowledge on various aspects of their academic and lived experiences. Academic Twitter users, including students, supervisors, and institutions, should actively promote positive and constructive online engagement. This can be achieved by engaging with helpful resources such as handles like @PHDcomics, @PhDVoice and @ThePhDPlace; and hashtags like #phdchat and #AcademicTwitter which provide empathetic responses to venting tweets and



offering supportive advice to those in need.

2. Reporting systems: Institutions should establish formal channels for PhD students to express their grievances and concerns about toxic supervision. This could include anonymous reporting systems, regular feedback sessions with supervisors, or dedicated support groups where students can openly discuss their experiences, express their feelings without fear. Institutions could also actively monitor discussions on social media platforms like Twitter to identify emerging issues related to toxic supervision and respond to them proactively.
3. Comprehensive training and support programmes: Institutions should develop and implement comprehensive training and support programmes for both supervisors and PhD students to address identified issues and foster a healthier academic environment. Supervisory training should cover interpersonal skills, cultural awareness and sensitivity, healthy work-life balance and ethical supervision practices.
4. PhD support services: Support services for PhD students to cope with the challenges of toxic supervision, such as counselling and mentorship programmes should be encouraged. Peer support groups should be encouraged and resources to help students navigate institutional policies and procedures should be instituted. For instance, peer support networks could be promoted to provide opportunities for students to share their experiences, exchange advice, and collectively advocate for change.
5. Institutional accountability: Mechanisms for holding institutions accountable should be established for addressing toxic supervision. The tasks should include implementing clear policies and procedures for handling complaints, providing avenues for anonymous reporting, and ensuring transparency in addressing reported incidents.
6. Synergetic Approach: In the planning and management of supervision, aspects of supervisory practices that must be considered to improve the supervisor-supervisee relationship include relationship dynamics, mental health, the support system and academic culture. The interventions should encompass a synergistic approach involving the institution, supervisors and students.

## References

Al Makhamreh, M., & Stockley, D. (2020). Mentorship and well-being: Examining doctoral students' lived experiences in doctoral supervision context. *International Journal of Mentoring and Coaching in Education*, 9(1), 1-20. <https://doi.org/10.1108/IJMCE-02-2019-0013>

Alam, F., Alam, Q., & Rasul, M. G. (2013). A pilot study on postgraduate supervision. *Procedia Engineering*, 56, 875-881. <https://doi.org/10.1016/j.proeng.2013.03.210>

Andrew, M. B. (2023). Neo-neoliberalist capitalism, intensification by stealth and campus real estate in the modern university in Aotearoa/New Zealand. *Journal of Applied Learning & Teaching*, 6(2), 393-401. <https://doi.org/10.37074/jalt.2023.6.2.16>

Barrero, A. F. (2023, February 17). J. L. Austin and John Searle on speech act theory. *The Collector*. <https://www.thecollector.com/speech-act-theory-austin-and-searle/>

Bogelund, P. (2015). How supervisors perceive PhD supervision and how they practice it. *International Journal of Doctoral Studies*, 10, 39-55. <https://doi.org/10.28945/2096>

Brookfield, S. D., Rudolph, J., & Tan, S. (2022). Powerful teaching, the paradox of empowerment and the powers of Foucault: An interview with Professor Stephen Brookfield. *Journal of Applied Learning & Teaching*, 5(1), 131-145 <https://doi.org/10.37074/jalt.2022.5.12>

Chen, K., Yang, S., & Duan, Z. (2021). Twitter as research data: Tools, costs, skillsets, and lessons learned. *Politics and the Life Sciences*, 41(1), 114-130. <https://doi.org/10.1017/pls.2021.19>

Chugh, R., Macht, S., & Harreveld, B. (2022). Supervisory feedback to postgraduate research students: A literature review. *Assessment & Evaluation in Higher Education*, 47(5), 683-697. <https://doi.org/10.1080/02602938.2021.1955241>

Costa, K. A. (2018). *Systematic review of challenges in research supervision at South African universities*. <https://doi.org/10.20944/preprints201812.0305.v1>

Dericks, G., Thompson, E., Roberts, M., & Phua, F. (2019). Determinants of PhD student satisfaction: The roles of supervisor, department, and peer attributes. *Assessment & Evaluation in Higher Education*, 44(7), 1053-1068. <https://doi.org/10.1080/02602938.2019.1570484>

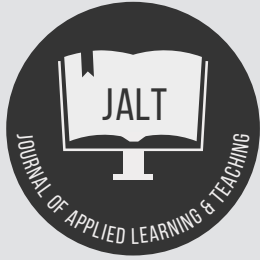
Ferreira, J. T. (2021). Reflecting on the viability of Twitter as tool in the postgraduate supervision process. *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.705451>

Foucault, M. (2000). *Power: Essential works of Foucault, 1954-84*. Penguin.

Gorup, M., & Laufer, M. (2020). *More than a case of a few bad apples: When relationships between supervisors and doctoral researchers go wrong*. Elephant in the Lab. [https://zenodo.org/records/4213175/files/EitL\\_Gorup%26Laufer.pdf?download=1](https://zenodo.org/records/4213175/files/EitL_Gorup%26Laufer.pdf?download=1)

Gruzdev, I., Terentev, E., & Dzhafarova, Z. (2020). Superhero or hands-off supervisor? An empirical categorization of PhD supervision styles and student satisfaction in Russian universities. *Higher Education*, 79(5), 773-789. <https://doi.org/10.1007/s10734-019-00437-w>

- Haven, T., Bouter, L., Mennen, L., & Tjldink, J. (2022). Superb supervision: A pilot study on training supervisors to convey responsible research practices to their PhD candidates. *Accountability in Research*. <https://doi.org/10.1080/08989621.2022.2071153>
- Hazell, C. M., Chapman, L., Valeix, S. F., Roberts, P., Niven, J. E., & Berry, C. (2020). Understanding the mental health of doctoral researchers: A mixed methods systematic review with meta-analysis and meta-synthesis. *Systematic Reviews*, 9, 197. <https://doi.org/10.1186/s13643-020-01443-1>
- Jabre, L., Bannon, C., McCain, J. S. P., & Eglit, Y. (2021). Ten simple rules for choosing a PhD supervisor. *PLoS Computational Biology*, 17(9), e1009330. <https://doi.org/10.1371/journal.pcbi.1009330>
- Kis, A., Tur, E. M., Lakens, D., Vaesen, K., & Houkes, W. (2022). Leaving academia: PhD attrition and unhealthy research environments. *PLoS One*, 17(10), e0274976. <https://doi.org/10.1371/journal.pone.0274976>
- Liu, L., & Woo, B. K. P. (2021). Twitter as a mental health support system for students and professionals in the medical field. *JMIR Medical Education*, 7(1), e17598. <https://doi.org/10.2196/2F17598>
- Malik, A., Heyman-Schrum, C., & Johri, A. (2019). Use of Twitter across educational settings: A review of the literature. *International Journal of Educational Technology in Higher Education*, 16, Article 36. <https://doi.org/10.1186/s41239-019-0166-x>
- Masek, A., & Alias, M. (2020). A review of effective doctoral supervision: What is it and how can we achieve it? *Universal Journal of Educational Research*, 8(6), 2493-2500. <http://dx.doi.org/10.13189/ujer.2020.080633>
- Moore, R., Rudling, E., Kunda, M., & Robin, S. (2021). Supporting casual teaching staff in the Australian neoliberal university: A collaborative approach. *Journal of Applied Learning and Teaching*, 4(2), 54-67. <https://doi.org/10.37074/jalt.2021.4.2.8>
- Muraranaza, C., Mtshali, N., & Bvumbwe, T. (2020). Challenges in postgraduate research supervision in nursing education: Integrative review. *Nurse Education Today*, 89, 104376. <https://doi.org/10.1016/j.nedt.2020.104376>
- Muthanna, A. A., & Alduais, N. A. (2021). A thematic review on research integrity and research supervision: Relationships, crises and critical messages. *Journal of Academic Ethics* 19(1), 95-113. <https://doi.org/10.1007/s10805-020-09374-w>
- O'Hara, C., & Cook, J. M. (2018). Doctoral-level counseling students' experiences of social class microaggressions. *Counselor Education and Supervision*, 57(4), 255-270. <https://doi.org/10.1002/ceas.12115>
- Payne, L. (2021). *One in four PhD students drop out*. Palatinat. <https://www.palatinat.org.uk/one-in-four-phd-students-drop-out/>
- Priyadarshini, M., Gurnam, K. S., Teoh, S. H., Geethanjali, N., & Chan, Y. F. (2022). Key factors influencing graduation on time among postgraduate students: A PLS-SEM approach. *Asian Journal of University Education*, 18(1), 51-64. <http://dx.doi.org/10.24191/ajue.v18i1.17169>
- Searle, J. (1979). *Expression and meaning*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511609213>
- Shahnawaz, M. G., & Siddiqi, N. (2022). Examining toxic supervision in higher education in India. *Higher Education Evaluation and Development*, 17(1). <https://www.emerald.com/insight/content/doi/10.1108/HEED-06-2021-0047/full/html>
- Tepper, B. J., Simon, L., & Park, H. M. (2017). Abusive supervision. *Annual Review of Organizational Psychology and Organizational Behavior*, 4(1), 123-152. <http://dx.doi.org/10.1146/annurev-orgpsych-041015-062539>
- van Rooij, E., Fokkens-Bruinsma, M., & Jansen, E. (2021). Factors that influence PhD candidates' success: The importance of PhD project characteristics. *Studies in Continuing Education*, 43(1), 48-67. <https://doi.org/10.1080/0158037X.2019.1652158>
- Young, S. N., VanWye, W. R., Schafer, M. A., Robertson, T. A., & Poore, A. V. (2019). Factors affecting PhD student success. *International Journal of Exercise Science*, 12(1), 34-45. <https://digitalcommons.wku.edu/ijes/vol12/iss1/3>



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## The challenge of making relationships central in online cultural safety education

---

Robyn Moore <sup>A</sup>	A	<i>Lecturer, School of Social Sciences, University of Tasmania</i>
Kim McLeod <sup>B</sup>	B	<i>Senior Lecturer, School of Social Sciences, University of Tasmania</i>
Natalia Maystorovich Chulio <sup>C</sup>	C	<i>Honorary Research Fellow, University of Sydney</i>

---

### Keywords

Allied health education;  
cultural safety education;  
culturally responsive pedagogy;  
emotional labour;  
learning;  
teaching.

---

### Correspondence

robyn.moore@utas.edu.au <sup>A</sup>

---

### Article Info

Received 8 August 2023  
Received in revised form 4 October 2023  
Accepted 5 January 2024  
Available online 11 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.7>

### Abstract

Cultural safety education entails the pedagogical strategy of taking students on a journey of discovery. This requires sustained openness to uncertainty, which can present myriad challenges for students and teachers. Learning about cultural safety is enabled when respectful, productive relationships characterise classrooms. In this paper, we report on the collaborative, reflective observations made by a group of university educators. We discuss educators' efforts to facilitate positive relatedness in online classrooms compared with their experiences in shared physical space (SPS) classrooms. We found that online environments enable and constrain relational possibilities in ways that differ from SPS classrooms and which escalate educators' emotional labour. Our findings highlight the significant role material/technological affordances of learning and teaching environments play in shaping relational possibilities. We argue that considering how the proximate materials and technologies in classrooms mediate relationship-building and connection needs to be factored into curriculum design and teaching practice. We propose drawing on culturally responsive pedagogies at the outset of cultural safety education design across SPS and online environments to prioritise relationship-building in ways that both enable students' learning and support educators' emotional labour.

## Introduction

Health professions in Australia increasingly expect graduates to enact culturally safe health care. A culturally safe environment affirms the manifold aspects of a person's lived experience (Bennett & Gates, 2019) and ensures no assault, challenge, or denial of any aspect of a person's identity (Williams, 1999). Culturally safe health care with Indigenous Australians requires health workers to support Indigenous Australians' sovereignty and demands. This requires "the ongoing critical reflection of health practitioner knowledge, skills, attitudes, practising behaviours and power differentials in delivering safe, accessible and responsive healthcare free of racism" (Australian Health Practitioner Regulation Agency, 2021, p. 9).

Cultural safety education requires a transformative learning process that intersects with students' and teachers' lives by exploring the self and one's beliefs, attitudes, and values. Students are invited to step into a lifelong process of considering what it means to be culturally safe in interaction with others (Best, 2018). Learning about cultural safety engages teachers' and students' emotions and bodies (Deckman & Ohito, 2020; James et al., 2022; Leonardo & Zembylas, 2013). Vulnerabilities and affective responses feature in the classrooms and require attention and care (Hollinsworth, 2016). Effective cultural safety education requires teachers to take a shared co-learning stance with students (McLeod, Moore et al., 2020), consistent with asking students to consider that we are "always implicated in each other's lives" (Baltra-Ulloa, 2018, p. 135). Teachers must 'walk the talk' and aim to model and enact ways of relating that cultivate the "atmosphere of openness, approachability, fairness, and safety" that Phan et al. (2009, p. 328) indicate is necessary when teaching in this space. Importantly, student-educator relationship building is essential to ensuring the classroom itself is culturally safe for students from marginalised groups, including Aboriginal and Torres Strait Islander students (Fernando & Bennett, 2019).

Cultural safety education aims to enable students to create new frameworks of thinking in which differences are legitimised. McDermott (2012, p. 15) describes the pedagogical strategy of taking students on a journey of discovery, which requires a sustained openness to uncertainty. The challenges for students, when presented with this learning opportunity, can manifest as resistance (Denis, 2011; Gatwiri, 2018; Hollinsworth, 2016). Cultural safety education is characterised by the emotional labour involved in facilitating and participating in these "hard conversations" (Sjorberg & McDermott, 2016, p. 29). Teachers must "anticipate the discomfit of disruption" (McLeod, Thakchoe et al., 2020, p. 187) associated with unlearning processes and work productively with this resistance in the classroom (Sjorberg & McDermott, 2016).

The quality of relationships between students and teachers in the cultural safety education classroom is critical to ensuring a productive learning process (McGill et al., 2021). Relatedness is defined as an individual's experience of relationships with others where there is a "sharing of meaningful feelings including warmth and affection in human contact" (Hagerty et al., 1993, p. 292). Respectful

connections between students and teachers generate a learning environment where students can create new understandings and insights from being ontologically disturbed (Ohito & Oyler, 2017). As McDermott (2012, p. 15) observes, "good cultural-safety education generates disquiet but makes the uncomfortable comfortable enough, through sensitive classroom facilitation in a mutually respectful environment". For Gill (2022), this objective can be achieved by establishing 'brave spaces' in which discomfort is explicitly acknowledged while authenticity and vulnerability are facilitated. As Pawlowski (2018, p. 63) asserts, "[b]rave space assumes that tension, conflict, and risk are at the heart of the cognitive and personal transformation". Within 'brave spaces', students are encouraged to rise to the challenges of genuine dialogue (Hole & De Luz, 2022).

Challenges associated with establishing the relationships that are imperative to effective cultural safety education can be amplified in online classrooms. As Powell et al. (2021) observe, "there remain critical questions around how best to ensure student engagement within the online environment" (p.1). The debate about how students engage in online classrooms also identifies its potential for supporting students' learning and transformation. For example, Hodges et al. (2020) suggests that online teaching can produce sound pedagogical outcomes. Similarly, Canty et al. (2020) state that the increasing range of online technologies can provide "high-quality distance learning that is engaging, interactive and increasingly personalised" (p. 3). Social interaction where students can share their values and interests has been identified as essential to student learning in online spaces (Alqurashi, 2019; Tang & Tsui, 2018) and there are calls for more teaching strategies to facilitate social interaction online (Baber, 2022).

This project contributes to the literature about cultural safety education by exploring how the online space mediates teachers' experiences of forming relationships with students and creating the 'brave spaces' that facilitate productive learning and teaching environments. We contrast online and SPS classrooms with the aim of contributing to the development of knowledge about best practice in relation to cultural safety education. The rapid shift to online learning in tertiary institutions due to COVID-19 (Crawford et al., 2020; Hodges et al., 2020) provides a unique opportunity to investigate these claims by comparing pre- and post-online teaching and learning experiences. This project resonates with other explorations of replicating SPS teaching online during the unique conditions of COVID-19 (Dinh & Nguyen, 2020). To this end, we asked:

- How do teachers experience the delivery of cultural safety education online compared to SPS classrooms?
- What are the strengths and weaknesses of online delivery for cultural safety education?

## Methods

This article examines the experiences of five educators from three campuses at a regional university. Two participants in the project are also members of the research team (authors one and three). All educators teaching cultural safety to allied health students at the university at the time (n=12) were invited to participate. At the time of the study, all participants were casual employees, either PhD students or early career researchers. Educators who agreed to participate submitted consent forms. This study was approved by the University's Human Research Ethics Committee.

Participants were responsible for delivering six two-hour online cultural safety workshops as a mandatory course component for allied health students. The workshops covered race, ethnicity, religion, disability, gender and sexuality. Following this content, students undertake a discrete unit about culturally safe health care and Aboriginal and Torres Strait Islander peoples. All educators had previously facilitated the workshops in SPS classrooms and were required to abruptly shift their teaching to online in response to COVID-19.

At three points during teaching delivery, participants reflected on their experiences. Initially, participants submitted written reflections to a shared Microsoft Teams folder. After the submission of each set of written reflections, a collaborative reflective conversation (CRC) was facilitated by Author Two, who was not a participant. CRCs were audio recorded with consent and transcribed verbatim. Thus, the data for this project comprised 15 written reflections and three transcribed CRCs.

Although our sample size is relatively small (n=5), the iterative nature of data collection supports in-depth exploration of educator experiences. The sample size also meets the requisite criteria outlined in Malterud et al. (2015) and Morse (2000) in relation to aims and scope of the study, study design, analysis strategy and quality of the data. Our sample size allows for diverse experiences to be explored while also enabling a thorough thematic analysis of the qualitative data (Boulton & Hammersley, 2006).

In line with the exploratory nature of this project, individual educator reflections were not guided by instructions or prompts beyond the project's information sheet. The aim was to allow topics not previously discussed in the literature to emerge. Nevertheless, to ensure that our research question was addressed in sufficient detail, the CRC facilitator guided participants to explore their experiences in relation to the specificity of cultural safety education, in particular, the challenges and rewards of teaching online.

This mixed-methods project is collaborative in design and analysis. Collaborative team research is increasingly recognised as more effective and productive as it tends to achieve greater outcomes than research conducted in isolation (Kelly et al., 2020). Mixed-methods research is becoming accepted as the third research approach (Johnson et al., 2007; Terrell, 2012; Molina, 2016). By combining individual written reflections with collaborative reflective conversations, we aimed to increase the depth with which

the research questions were explored.

Individual reflections were chosen for this project due to the wealth of literature that describes journaling as a process that supports professional development for teachers. For example, journaling has been demonstrated to support teachers to increase their understanding, connect with others and pose questions (Alterio, 2004; Göker, 2016). Sharing their reflections with other participants and contributing to CRCs supported the development of collegial relationships as participants jointly explored classroom dynamics and teaching experiences (MacPherson, 2010). In this way, the CRCs were not solely focused on data collection, but also designed to cultivate a community of practice (Sumer et al., 2021). In line with our aim of building collegial relationships, reflections shared to the Teams folder were not anonymised. However, participant extracts in this paper have been anonymised.

Grounded theory is a qualitative methodology that emphasises a systematic inductive approach to data collection and analysis focusing on building theory from data rather than hypotheses (Corbin & Strauss, 1990). Grounded theory was chosen because its inductive principles align with the exploratory aims of this research, allowing us to generate new theories about the experiences of teaching cultural safety online, where little previous research exists. Following the grounded theory method, data were analysed first by in vivo coding (Saldaña, 2014, p. 590). All researchers independently coded the data using the participants' actual language. After sharing our individual codes, we met online to collaboratively develop a coding framework by identifying higher-order codes. We then jointly categorised the data according to this framework.

While time consuming, the collaborative nature of data analysis allows for individual readings of the data to be critically examined. The limited size of our research sample rendered this method of data analysis feasible. Our coding framework comprises the following higher-order codes: 1. Mechanics of the online space; 2. Interactions—peer-to-peer and student-teacher; 3. Teacher observations of student learning; 4. Teacher identity; and 5. Teaching strategies. These codes were further refined during the writing phase to produce our final two themes. Our first theme discusses the implications of online teaching environments lacking many of the often taken-for-granted benefits of SPS classrooms. Our second theme explores how these affordances of online environments affect student-teacher relationships. These themes are explored in turn below.

## Analysis and discussion

### *The affordances of online space enable and constrain relationship-building*

Teaching environments have traditionally been predicated on people being physically present in a classroom. The interactions that take place in the classroom offer both the teacher and students the capacity to connect not only through words, but also through body language. Converting the classroom to an online space has implications for the

interpretation of non-verbal responses, especially if cameras are not used. This can be crucial for cultural safety learning, which engages teachers' and students' emotions and bodies (Deckman & Ohito, 2020; James et al., 2022; Leonardo & Zembylas, 2013). On the other hand, online classrooms may hold increased potential for collaboration between teachers and students by undermining traditional power dynamics. The relative anonymity of online classrooms may increase student sharing if students feel less vulnerable in comparison to SPS classrooms (Malbon, 2022). For example, Stacy described her surprise at students' willingness to share personal information in relation to their own identity and how the online space created affordances for collaboration and sharing of experiences.

The greater anonymity that students have online may be one of the reasons they seem to have been more willing to give their opinions and take part in a discussion although, as with face-to-face teaching, there were still a few who said nothing. (Stacy, RJ [Reflective Journal]1)

Stacy recognises that, for some students, the online space created possibilities of relatedness and sharing that were sometimes missing in the SPS classroom. In her experience, the online space, particularly the written chat function, provided students with anonymity and democratised the learning experience. Other participants commented that, rather than a few 'vocal' students dominating discussions as often occurred in SPS classrooms, the conversation was more evenly shared in the online environment.

In person when I was teaching, there were some people who would really carry the conversation...but I feel like there is a lot more shared, even though it is a bit like getting a text message and there's not much detail. (Mackenzie, CRC [Collaborative Reflective Conversation]1)

Educators attributed students' increased willingness to participate to their use of the written chat function of online classrooms, which was by far the most prevalent method of student interaction. The chat function established a more democratic and safe space for students to contribute their experiences and ideas than might have been possible in the SPS.

Largely the class used the [written] chat function offering short and concise responses to the questions. (Octavia, RJ2)

Most students seemed more comfortable commenting on, and exploring content, via the [written] chat function...I feel like this function, allowed students—who might otherwise feel uncomfortable commenting—feel safe, and able to make points and respond to questions. (Mackenzie, RJ1)

The written chat function also offered opportunities for peer learning as students engaged with each other, often "asking relevant questions" (Selmah, CRC1). For the most part, however, student reliance on short, perfunctory written

responses left educators questioning student engagement (Chen et al., 2020) and their ability to 'gauge the room' and adjust their delivery accordingly. Crucially, short written responses did not allow for more nuanced conversations, restricting the capacity of teachers and students to develop their relationships. Our participants' observations about limited engagement through short written responses connect with the findings of Mulrooney and Kelly's (2020) study, which stresses the critical role of relationships in virtual learning. They affirm that building strong teacher-student relationships is pivotal for fostering a sense of belonging and enhancing overall learning experiences while observing that online environments present challenges to developing these relationships.

As research participants reflected on teaching online, fundamental questions frequently arose about how to assess student engagement, and what constitutes a good learning experience for students. In this section we discuss implications arising from the lack of embodied responses, the severing of happenstance interactions and the benefits of dedicated in-person study spaces.

Effective classroom management and pedagogical approaches contribute to engagement, but this is predicated on student-teacher relationships, the teacher's capacity to read the room and their ability to create 'brave spaces' (Hole & De Luz, 2022). These features become challenging when classes move from SPS to online (Lonie & Andrews, 2009). Participants explained that students generally did not use their cameras and microphones to engage in the online space and this impacted the capacity for the teacher to establish relationships in the classroom. Kedraka and Kaltsdis (2020) similarly discovered that students exhibited a marked preference for written communication when posing questions, displaying notable reluctance to use microphones and cameras. Interestingly, these same students noted the loss of interaction and connectedness in the virtual space affected their learning experiences.

Stacy (RJ1) reflected that it was "strange and disarming" to speak into the webcam of her computer and "not seeing faces or hearing voices". The silence experienced by teachers made the process difficult. As Mackenzie (CRC3) described, "there was just this silence from the group", making it difficult to establish cultures of relatedness so that students felt safe sharing their experiences and beliefs.

For me, the main classroom was like talking to the empty room I was sitting in. I felt pressured to talk to get some sort of discussion going. (Ewan, RJ1)

Research participants questioned how they received or read engagement in students when the embodied demeanour or facial and bodily expressions were absent, with their only clues being limited comments and questions shared by audio or typed in the chat. Chen et al. (2020, pp. 224-225) observe that the loss of body language and non-verbal cues in online learning environments force teaching staff to consistently check in on the student's comprehension of material. This was also true for participants in our study. For example, Stacy (CRC2) stated that she consistently asked the students for feedback on how they were progressing through the material.

It was only in the process of teaching online that participants realised the value of embodied responses, which they had previously taken for granted. They missed the common cues that point to “how students [are] receiving and responding to the material” (Stacy, RJ1). As Mackenzie noted, students’ reliance on participating via written chat meant that it was difficult to know if students understood the task at hand or merely gave responses that they believed were sought by the educator:

I don’t know if they got it, or if they were just saying what they wanted me to hear...but it was a lot more difficult to tell. (Mackenzie, CRC1)

This restricted level of engagement and lack of input from students was perceived by educators as a lack of engagement (Chen et al., 2020), which led them to experiencing teaching as one-sided with feelings of being surveilled. Mackenzie described the experience as similar to an interview and Octavia likened it to Bentham’s panopticon.

Sometimes when people just say things like, “Yes, I agree” or “I feel the same as such and such”, it kind of feels like I’m almost interviewing them as you would a research participant...as opposed to just discussing ideas or helping thinking. (Mackenzie, CRC2)

[Teaching online] reminded me of Jeremy Bentham’s idea of the panopticon...it was this feeling of being constantly under surveillance. (Octavia, CRC1)

Research participants observed that, in online classrooms, some forms of contact and communication between students, and students and teachers, no longer happened. For example, participants noted the absence of the informal conversations between students that are possible when students move in and out of the classroom together. They noted how the incidental learning that happens when students discuss unit content as part of their social connections with other students is missing in online classrooms.

A further benefit of SPS classrooms is that they support students to be present both physically and mentally with, ideally, minimal distractions from classroom activities. In contrast, students in virtual classrooms may be joining from busy home or work environments rather than a quiet study space. Anecdotal evidence reveals that online students may be distracted and passively listening or embarrassed to share their personal space on camera (Stafford, 2020, pp. 150-151). As Thatsara et al. (2020, p. 44) note, the lack of a “proper study” environment at home creates difficulties for student engagement. The distractions typical of home environments may have been exacerbated during our project due to COVID-19 lockdowns in which all family members were home and caring responsibilities were often increased. In addition to making learning difficult for students, this can also create disruptions for others in the virtual space.

When breakout groups were closed and students returned, their microphones were still on and you could hear the noise and that others were in the same space during class. I reflected on how difficult

it must be for students—to have other things taking place in the background. (Octavia, RJ1)

With students no longer in physical proximity to each other, a key dynamic the research participants observed in SPS classrooms, where students readily form groups and generate a commentary together to seize or resist the learning opportunity, was disrupted. Research participants shared how students can aggregate to create what Selmah described as a “negative downward spiral or you can have a really positive spiral” (CRC3). This occurs when one or a few students carry a conversation and propel other students into a negative or positive and robust discussion.

It’s easier (for me) when students absent themselves rather than having to deal with active resistance, silence, refusal to engage while present in person. Is it better for the students though? Do we hope to get through to the ‘resisters’ by the discussion, activities, peer modelling? If the students are present then, in theory, there’s a possibility of them changing their perspective. (Selmah, RJ1)

This points to the importance of establishing strategies to foster relationships that enable teachers to critically engage with negative discussions and create ‘teachable moments’, something not possible if there is limited interaction between the cohort and teacher. Research participants discussed how their experiences of online classrooms prevented the conditions they know are required to facilitate critical discussion and the necessary ‘ontological disturbance’ in students. As Ewan observes, fostering student-teacher relationships online can be challenging due to the nature of virtual learning platforms.

I want to have difficult, tough conversations that are going to ontologically disturb them, but you need to bring people close to that for that and you can’t in this environment. (Ewan, CRC3)

Overall, the educators participating in our research project felt that online classrooms lacked many of the often taken-for-granted benefits of SPS classrooms—embodied responses, happenstance interactions and dedicated study spaces. This left our participants feeling surveilled and increased the emotional labour of teaching. In the next section, we discuss the implications of the online environment for participants’ sense of developing the relationships that are fundamental to cultural safety teaching.

### ***Constrained relationship-building in online spaces escalates teachers’ emotional labour***

The quality of relationships between students and teachers in the cultural safety education classroom is critical to ensuring a productive learning process (McGill et al., 2021). These relationships demand a mutually respectful environment (McDermott, 2012) in which educators model ways of relating that cultivate openness and co-learning. This orientation to teaching necessitates the emotional labour (Hochschild, 1983) of managing and regulating emotions—a

widely observed part of teachers' work in SPS classrooms (Constanti & Gibbs, 2004; Dismore et al., 2019). Research participants in this project found that additional dimensions of emotional labour were associated with attempting to create relatedness in online environments. For example, some research participants described how they mitigated the perceived lack of engagement by students, by becoming much more performative to encourage interactions and participation in the virtual classroom. Selmah commented that her "performance as a teacher felt quite contrived" as she sought to do "whatever [she] could, to keep the students interested and engaged". Ewan shared similar sentiments:

I was performing a bit more than I usually would. I know my teaching style is a bit performative, as a way to loosen people up and welcome them in, but it felt really forced. (Ewan, RJ1)

This experience is supported by Chen et al. (2020, p. 230) who found that online platforms required teachers to "rehearse their performance...more like 'show time'". However, performativity is not conducive to developing 'brave spaces' (Hole & De Luz, 2022) to enable the relationships necessary for effective cultural safety education. When educators model performativity rather than vulnerability, openness and co-learning, the conditions for facilitating the "hard conversations" (Sjorberg & McDermott, 2016, p. 29) that produce transformative learning outcomes are not created. These limitations are not conducive to supporting students to critically reflect on their own beliefs and attitudes.

The additional demands of the virtual classroom (Lavine et al., 2012) necessitate different kinds of emotional labour, which left the participants in our study feeling drained. This finding resonates with Nyanjom and Naylor (2021), who assert that the emotional labour of online teaching can have negative impacts on teachers' well-being. This feeling was augmented by technological difficulties, which disrupted educators' focus on cultivating a productive classroom environment for cultural safety education. Echoing the experiences of many teachers forced to switch to online delivery during the COVID-19 pandemic (Martin, 2020), participants in our study noted issues with technology and bandwidth. The proliferation of online teaching globally (Arday, 2022; Crawford et al., 2020; Mulrooney & Kelly, 2020) has prompted discussions about the importance of technological tools (Baran et al., 2011; Eri et al., 2021; Kaqinari et al., 2021; Sumer et al., 2021) to establish virtual learning spaces through chat rooms, video meeting spaces and interactive whiteboards (Crawford-Ferre & Wiest, 2012; Major, 2015; Montelongo & Eaton, 2019). However, this is predicated on having access to required equipment and bandwidth, and for all the systems to operate efficiently. While not a focus of this research project, it is important to note that the rapid shift to online learning during the COVID-19 pandemic showed that student access to equipment or digital capacity cannot be assumed (Mshigeni et al., 2020). As noted by Arday (2022), the pandemic exacerbated inequalities, laying bare the flaws in the system.

Educators' attempts to create effective learning environments while managing technological issues often proved challenging. Bower et al. (2014) found that the technical

difficulties required teachers to make snap decisions under pressure. We noticed similarities with how participants in our study described their experiences with internet connectivity issues.

Soon after the introductions and icebreaker activity, however, things started to go awry...my connectivity was low...I was dropping out occasionally. (Stacy, RJ1) Whether it was my Internet, my laptop, or the MyLO system, halfway through the final workshop—when I attempted to create break-out groups—the site froze. (Mackenzie, RJ1)

During the tutorial, I experienced my own issues with the bandwidth... forcing me to switch from WIFI to a mobile hotspot. (Octavia RJ1)

In the second workshop, I had 11 students and they consistently dropped out of the session. At one point when I was providing instructions before putting them in breakout groups, all 11 students dropped out of the session. (Octavia, RJ2)

Struggling with bandwidth and having the teacher and/or the students drop out of the virtual classroom can be very disruptive and stressful, taking away from the time set for facilitating the learning process (Martin, 2020). The need to manage the technology added an additional layer of complexity, which at times distracted participants in this project from engaging and relating to the cohort in ways that are conducive for cultural safety learning. Participants reflected on the different experiences between online and face-to-face teaching.

I suddenly missed the classroom experience. Whilst I have encountered technical difficulties with the projector or audio-visual equipment in face-to-face teaching, nothing had ever been so disruptive, and help was at hand. (Stacy, RJ1)

It is futile to focus on cultivating supportive relationships while negotiating connectivity issues. The need to manage technology impacted participants' capacity to create an optimal classroom environment, leaving them feeling exhausted and disillusioned with the experience. This study shows that technological issues made communication harder and increased participants' emotional labour (Kennedy et al., 2022).

Participants in this study emphasised the increased emotional labour associated with a sense of lost reciprocity in learning experiences with students. The most prevalent theme in educator reflections was student reliance on the written chat function in the online classroom and implications arising from this. Students' inability or unwillingness to engage via video and audio was often perceived as a lack of engagement and/or reciprocity (Chen et al., 2020). In combination with teachers' attempts to engage students, the perceived lack of reciprocity left teachers feeling "exhausted", "tired", "drained" or "pooped".



By the end of the workshops, I felt completely drained and exhausted after having to be switched on and engaging while my students were 'just there'. It is really hard to be talking to students and not getting anything back from them. The experience felt like being on radio or doing a podcast and not knowing how the audience is experiencing the learning experience. (Octavia, RJ1)

Across the teaching team, the perceived lack of student engagement left our participants struggling to feel like effective educators. Similarly, Stafford (2020, p. 151) reflects that lack of student engagement contributed to "unmet expectations and frustration for teachers" who were not accustomed to such student behaviours. In our study, while participants experienced "unmet expectations and frustration", this was combined with feeling "disheartened", "inadequate" and "dejected".

No one talked. And it was really hard work just getting them to even respond in the chat box...I just sort of sat there in silence and just waited until someone got so uncomfortable that you know that they'd say something. It was awful, yeah, really hard. (Stacy, CRC2)

And I feel like an absolute failure at the end. Like, I'm just thinking 'what am I doing wrong'? (Octavia, CRC2)

It was incredibly hard to get people to engage. (Mackenzie, RJ2)

Participants' comments above highlight the emotional labour involved in the teaching process and how this impacts their experiences. Similar observations about the impact of the rapid shift to online learning during COVID-19 on the well-being of educators has been found in other studies (Konstantinou & Miller, 2022; Kennedy et al., 2022; Nyanjom & Naylor, 2021). In our study, the ability to engage students in critical and reflective conversations was hampered by the technological limitations of attempting to replicate SPS workshops in an online environment. As Lee (CRC2) noted, "in this space, we're being set up to fail". This has implications for teachers as high levels of emotional labour are demanded in online teaching spaces in order to achieve any level of successful engagement with cultural safety education to prepare students for their future professions. Participants felt that the predominantly 'unembodied' nature of teaching online impeded the development of the relationality that is crucial to effective cultural safety education.

Participants' comments above demonstrate our research participants' commitment to their teaching and the effort they perform regardless of the learning and teaching environment. Participants' commitment to their teaching was also evident in the suggestions made for improving learning experiences for students:

- Provide more clarity and explanation about course aims, the kind of learning journey to expect and how this may differ from other learning experiences

in their degrees. Centre relationality and deep listening as core learning outcomes.

- Mandate attendance by designing assessment items that must be completed in class, alone or working collaboratively with peers. This will help to emphasise the importance of the classroom interaction and enable assessment of how students engage with peers. Develop peer feedback tasks.
- Consider students' own insights into their ongoing learning about their relational capacities for collaboration and their sensitivity to difference. Provide more feedback to students about their learning journey.
- Develop learning activities in which students reflect on what happens in the workshops.

These examples reveal the repertoires of practice that educators draw on to create culturally safe classrooms.

## Conclusion

In this project, the abrupt shift to online teaching provoked observations and questions about how online classrooms mediated students' engagement with the learning activities, and how this contrasted with participation in SPS classrooms. The research participants in this project had all previously facilitated similar learning activities in physical classrooms. This enabled a direct comparison of the effectiveness of the learning activities in virtual classrooms. The comparison between SPS and online teaching environments shows that the affordances of the environments themselves mediate how relationships are built. This study supports previous research about cultural safety education in SPS classrooms, arguing that relationships and connection are vital to good teaching practice. Participants felt that the predominantly 'unembodied' nature of teaching online impeded the development of the relationality that is crucial to cultural safety education. In facilitating cultural safety learning with healthcare students, a key element is the development of relatedness to create a 'brave space' (Hole & De Luz, 2022) for students to share their values, ideas and experiences (Bennett et al., 2022). The development of respectful student-teacher and student-student relationships provides an essential foundation for individuals to discuss their values, ideas and experiences regarding confronting topics such as race, gender and sexuality. This study shows educators need to actively and creatively work with the technological affordances to facilitate connection in online spaces between students, and between teachers and students. We found the affordances of online classrooms somewhat enabled, but mostly constrained relationship-building. In this case, SPS interactive workshops were replicated online with insufficient time to update materials or modify the design of learning activities. The changed material and technological affordances substantially impacted the potential for relatedness. This study points to the importance of factoring into curriculum design and teaching practice how the proximate materials and technologies mediate relationship-building and connection in the classroom.

We propose drawing on culturally responsive pedagogies at the outset of cultural safety education design across SPS and online environments to prioritise relationship-building in ways that enable students' learning. Culturally responsive pedagogy has emanated from colonial settler countries such as Canada, Australia, New Zealand, and the United States, arising from the civil rights movement. Culturally responsive pedagogy sits within the critical pedagogy tradition, draws on sociocultural learning (the notion that learning is socially mediated and relates to students' cultural experiences), and views learning through an anti-deficit lens (Morrison et al., 2019). Key pedagogical approaches include drawing on and speaking to the realities of students' lives and using the cultural knowledge, prior experiences, frames of reference, and performance styles of diverse students to make learning encounters more relevant and effective for them (Gay, 2013). Positive and meaningful relationships between students, teachers, and their families and communities, are considered critical to enacting culturally responsive pedagogies (Shevalier & McKenzie, 2012).

The question of how to establish relatedness in online learning is a live question in the growing body of scholarship that explores how culturally responsive pedagogies apply to online learning environments (Lawrence, 2020). For Montelongo & Eaton (2019, p. 42), this extends "the traditions of Paulo Freire, Bell Hooks, Gloria Anzaldúa, and other critical standpoint theorists to deconstruct power dynamics in online learning spaces, centre relational and dialogic praxis" and furthers the development of critical digital pedagogy. Bennett et al. (2022, p. 1660) indicate the importance of respectful relationships in online teaching and draw on culturally responsive and trauma-informed pedagogical approaches, with the specific aim "to create a democratic, collaborative, and reflexive space whereby students and educators can feel simultaneously supported in the diversity of their respective lived experiences and learning". Scholars who draw on culturally responsive pedagogies provide examples of innovative teaching strategies that enable connection and communication between students and teachers in online spaces. Strategies include collaborative group work; creating a welcoming virtual environment via an assignment for students to introduce each other; using synchronous online meeting spaces (Woodley et al., 2017); immersive video experiences; chats between students; creating a course song playlist; activities to model active listening (Montelongo & Eaton, 2019) and building class community through whole class communication (Lawrence, 2020). Culturally responsive pedagogies support educators in bringing sustained attention to enabling relationships between students, and students and teachers, in both SPS and online environments, suggesting ways that educators can work actively with the affordances of the online teaching spaces to support connection and communication. Drawing on culturally responsive pedagogies at the outset of cultural safety education design across SPS and online environments is, we suggest, also a way to support educators' emotional labour. Culturally responsive pedagogies value all relatedness in the classroom, including with/from the educator. Teaching educators about culturally responsive pedagogies is a way of making emotional labour explicit in teaching practice, as something that can actively be engaged within the classroom for the purposes of

relationship-building. It is important to support this critical dimension of educators' work with professional and collegial learning which shares coping strategies (Nyanjom & Naylor, 2021) and enables links between emotions and beliefs, and professional practice (Fu & Clarke, 2023). Also critical is the institutional recognition, acknowledgment, and support for the emotional labour of online educators (Konstantinou & Miller, 2022; Nyanjom & Naylor, 2021). Given the high proportion of university teaching undertaken by casual staff, this requires consideration of how best to support casual educators (Moore et al., 2021). As Kennedy et al. (2022, p. 30) observes, "since emotional labour is often borne by the least privileged sections of the university workforce, this study uncovers uncomfortable questions about the persistence of systemic problems causing staff inequalities that cannot afford to be ignored."

We acknowledge the limitations of our study, particularly regarding the scope of the paper, which primarily focused on the teacher experience during the pandemic. Initially, the study sought to contrast online and SPS classrooms with the aim of contributing to the development of knowledge about best practices in cultural safety education. However, during the analysis phase it became clear that gaining insight into the student experience would have enhanced the findings by providing the research team with a different perspective. However, our study aligns and complements the findings of Mulrooney and Kelly (2020), who examined the student and teacher experience in relation to belonging in the United Kingdom. They too recognised the constraints of the online space for academic engagement caused by socio-economic conditions and the lockdown experience. Moving forward, future research should aim to investigate the experiences of teachers and students to attain a more comprehensive and well-rounded understanding. Additionally, this inclusion of student perspectives will enrich our findings and enable us to identify key areas that require attention and improvement in the realm of online education. By doing so, we can generate insights that have meaningful and actionable implications for cultural safety education.

## Acknowledgement

The authors wish to sincerely thank our colleagues who generously shared their time and experiences.

## References

- Alqurashi, E. (2019). Predicting student satisfaction and perceived learning within online learning environments. *Distance Education, 40*(1), 133-148. <https://doi.org/10.1080/01587919.2018.1553562>
- Alterio, M. (2004). Collaborative journalling as a professional development tool. *Journal of Further and Higher Education, 28*(3), 321-332. <https://doi.org/10.1080/0309877042000241788>
- Australian Health Practitioner Regulation Agency. (2021). *The national scheme's Aboriginal and Torres Strait Islander health and cultural safety strategy 2020-2025*. <https://www.>

- ahpra.gov.au/About-AHPRA/Aboriginal-and-Torres-Strait-Islander-Health-Strategy.aspx
- Arday, J. (2022). Covid-19 and higher education: The times they are a'changing. *Educational Review*, 74(3), 365-377. <https://doi.org/10.1080/00131911.2022.2076462>
- Baber, H. (2022). Social interaction and effectiveness of the online learning—A moderating role of maintaining social distance during the pandemic COVID-19. *Asian Education and Development Studies*, 11(1), 159-171. <https://doi.org/10.1108/AEDS-09-2020-0209>
- Balra-Ulloa, A. J. (2018). Speaking of care from the periphery: The politics of caring from the post-colonial margins. In B. Pease, A. Vreugdenhil & S. Stanford (Eds.), *Critical ethics of care in social work: Transforming the politics and practices of caring* (pp. 129-138). Routledge. <https://doi.org/10.4324/9781315399188>
- Baran, E., Correia, A. P., & Thompson, A. (2011). Transforming online teaching practice: Critical analysis of the literature on the roles and competencies of online teachers. *Distance Education*, 32(3), 421-439. <https://doi.org/10.1080/01587919.2011.610293>
- Bennett, B., & Gates, T. G. (2019). Teaching cultural humility for social workers serving LGBTQI Aboriginal communities in Australia. *Social Work Education*, 38(5), 604-617. <https://doi.org/10.1080/02615479.2019.1588872>
- Bennett, B., Ross, D., & Gates, T. G. (2022). Creating spatial, relational and cultural safety in online social work education during COVID-19. *Social Work Education*, 41(8), 1660-1668. <https://doi.org/10.1080/02615479.2021.1924664>
- Best, O. (2018). The cultural safety journey: An Aboriginal Australian nursing and midwifery context. In O. Best & B. Fredericks (Eds.), *Yatdjuligin: Aboriginal and Torres Strait Islander nursing and midwifery care* (2nd ed., pp. 46-66). Cambridge University Press.
- Boulton, D., & Hammersley M. (2006). Analysis of unstructured data. In R. Sapsford & V. Jupp (Eds.), *Data collection and analysis* (pp. 243-259). SAGE Publications Ltd. <https://doi.org/10.4135/9781849208802>
- Bower, M., Kenney, J., Dalgarno, B., Lee, M. J. W., & Kennedy, G. E. (2014). Patterns and principles for blended synchronous learning: Engaging remote and face-to-face learners in rich-media real-time collaborative activities. *Australasian Journal of Educational Technology*, 30(3), 261-272. <https://doi.org/10.14742/ajet.1697>
- Canty, A. J., Chase, J., Hingston, M., Greenwood, M., Mainsbridge, C. P., & Skalicky, J. (2020). Addressing student attrition within higher education online programmes through a collaborative community of practice. *Journal of Applied Learning & Teaching*, 3(S11), 1-12. <https://doi.org/10.37074/jalt.2020.3.s1.3>
- Chen, J. C., Dobison, T., & Kent, S. (2020). Lecturers' perceptions and experiences of Blackboard Collaborate as a distance learning and teaching tool via Open Universities Australia (OUA). *Open Learning: The Journal of Open, Distance and e-Learning*, 35(3), 222-235. <https://doi.org/10.1080/02680513.2019.1688654>
- Constanti, P., & Gibbs, P. (2004). Higher education teachers and emotional labour. *International Journal of Educational Management*, 18(4), 243-249. <https://doi.org/10.1108/09513540410538822>
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, 13, 3-21. <https://doi.org/10.1007/BF00988593>
- Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni P. A., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Learning & Teaching*, 3(1), 9-28. <https://doi.org/10.37074/jalt.2020.3.1.7>
- Crawford-Ferre, H. G., & Wiest, L. R. (2012). Effective online instruction in higher education. *Quarterly Review of Distance Education*, 13(1), 11-14. <https://link.gale.com/apps/doc/A297555182/AONE?u=usyd&sid=bookmark-AONE&xid=62f4b94d>
- Deckman, S., & Ohito, E. (2020). Stirring vulnerability, (un)certainly, and (dis)trust in humanizing research: Duoethnographically re-membering unsettling racialized encounters in social justice teacher education. *International Journal of Qualitative Studies in Education*, 33(10), 1058-1076. <https://doi.org/10.1080/09518398.2019.1706199>
- Denis, V. (2011). Silencing Aboriginal curricular content and perspectives through multiculturalism: 'There are other children here'. *Review of Education, Pedagogy, and Cultural Studies*, 33(4), 306-17. <http://dx.doi.org/10.1080/10714413.2011.597638>
- Dinh, L. P., & Nguyen, T. T. (2020). Pandemic, social distancing, and social work education: Students' satisfaction with online education in Vietnam. *Social Work Education*, 39(8), 1074-1083. <https://doi.org/10.1080/02615479.2020.1823365>
- Dismore, H., Turner, R., & Huang, R. (2019). Let me edutain you! Practices of student engagement employed by new lecturers. *Higher Education Research and Development*, 38(2), 235-249. <https://doi.org/10.1080/07294360.2018.1532984>
- Eri, R., Gudimetla, P., Star, S., Rowlands, J., Girgla, A., To, L., Li, F., Sochea, N., & Bindal, U. (2021). Digital resilience in higher education in response to COVID-19 pandemic: Student perceptions from Asia and Australia. *Journal of University Teaching & Learning Practice*, 18(5), 7-26. <https://doi.org/10.53761/1.18.5.7>
- Fernando, T., & Bennett, B. (2019). Creating a culturally safe space when teaching Aboriginal content in social work: A scoping review. *Australian Social Work*, 72(1), 47-61. <https://doi.org/10.1080/0312407X.2018.1518467>
- Fu, G., & Clarke, A. (2023). Connected by emotion: Teacher agency in an online science education course during

- COVID-19. *Journal of Research in Science Teaching*, 1-26. <https://doi.org/10.1002/tea.21886>
- Gatwiri, K. (2018). Leaning into the discomfort and embracing the disruption: A Freirean approach to (de)colonised social work teaching in Australia. *Whiteness and Education*, 3(2), 182-197. <https://doi.org/10.1080/23793406.2019.1573644>
- Gay, G. (2013). Teaching to and through cultural diversity. *Curriculum Inquiry*, 43(1), 48-70. <https://www.jstor.org/stable/23524357>
- Gill, K. R. (2022) Developing a syllabus policy on safety and comfort. *Double Helix*, 10.
- Göker, S. D. (2016). Use of reflective journals in development of teachers' leadership and teaching skills. *Universal Journal of Education Research*, 4(12A), 63-70. <https://doi.org/10.13189/ujer.2016.041309>
- Hagerty, B., Lynch Sauer, J., Patusky, K. L., & Bouwsema, M. (1993). An emerging theory of human relatedness. *Journal of Nursing Scholarship*, 25(4), 291-296. <https://doi.org/10.1111/j.1547-5069.1993.tb00262.x>
- Hochschild, A. R. (1983/2012). *The managed heart: Commercialization of human feeling*. University of California Press.
- Hole, B. V., & De Luz, M. (2022). An imaginary of radical hope: Developing brave space for classroom discussion. *Teaching Ethics*, 22(1), 83-96. <https://doi.org/10.5840/tej20221011119>
- Hollinsworth, D. (2016). Unsettling Australian settler supremacy: Combating resistance in university Aboriginal studies. *Race, Ethnicity and Education*, 19(2), 412-432. <http://doi.org/10.1080/13613324.2014.911166>
- James, T., Bond, K., Kumar, B., Tomlins, M., & Toth, G. (2022). We were all learning and doing our best: Investigating how enabling educators promoted student belonging in a time of significant complexity and unpredictability. *Journal of University Teaching & Learning Practice*, 19(4). <https://ro.uow.edu.au/jutlp/vol19/iss4/18>
- Johnson, R. B., Turner, L. A., & Onwuegbuzie, A. J. (2007). Toward a definition of mixed methods approach. *Journal of Mixed Methods Research*, 1(2), 112-133. <https://doi.org/10.1177/1558689806298224>
- Kaqinari, T., Makarova, E., Audran, J., Döring, A., Göbel, K., & Kern, D. (2021). The switch to online teaching during the first COVID-19 lockdown: A comparative study at four European universities. *Journal of University Teaching & Learning Practice*, 18(5). <https://doi.org/10.53761/1.18.5.10>
- Kedra, K., & Kaltsidis, C. (2020). Effects of the COVID-19 pandemic on university pedagogy: Student experiences and considerations. *European Journal of Education Studies*, 7(8), 17-30. <http://dx.doi.org/10.46827/ejes.v7i8.3176>
- Kelly, N., Doyle, J., & Parker, M. (2020). Methods for assessing higher education research team collaboration: Comparing research outputs and participant perceptions across four collaborative research teams. *Higher Education Research & Development*, 39(2), 215-229. <https://doi.org/10.1080/07294360.2019.1676199>
- Kennedy, E., Oliver, M., & Littlejohn, A. (2022). "You make yourself entirely available": Emotional labour in a caring approach to online teaching. *Italian Journal of Educational Technology*, 30(1), 30-48. <https://doi.org/10.17471/2499-4324/1237>
- Konstantinou, I., & Miller, E. (2022). Collegiality, emotional labour and compassion in the academy: An autoethnographic study of Covid-19. *Qualitative Research Journal*, 22(4), 593-605. <https://doi.org/10.1108/QRJ-07-2022-0095>
- Lavine, R., Greenburg, J., Chen, J., Kao, K., & Lin, Y. (2012). East meets west: A virtual international teacher education initiative between the US and Taiwan. *Journal of the International Society for Teacher Education*, 16(1), 35-45. <http://hdl.handle.net/20.500.11937/42934>
- Lawrence, A. (2020). Teaching as dialogue: Toward culturally responsive online pedagogy. *Journal of Online Learning Research*, 6(1), 5-33.
- Leonardo, Z., & Zembylas, M. (2013). Whiteness as technology of affect: Implications for educational praxis. *Equity & Excellence in Education*, 46(1), 150-165. <https://doi.org/10.1080/10665684.2013.750539>
- Lonie, A. L., & Andrews, T. (2009). Creating a rich learning environment for remote postgraduate learners [online]. *Education in Rural Australia*, 19(1), 3-13.
- MacPherson, S. (2010). Teachers' collaborative conversations about culture: Negotiating decision making in intercultural teaching. *Journal of Teacher Education*, 61(3), 271-286. <https://doi.org/10.1177/0022487109353032>
- Major, C. H. (2015). *Teaching online: A guide to theory, research, and practice*. JHU Press.
- Malbon, J. (2022). The intimacy of distance. In M. A. Peters, F. Rizvi, G. McCulloch, P. Gibbs, R. Gorur, M. Hong...& L. Misiaszek (Eds.), *Reimagining the new pedagogical possibilities for universities post-Covid-19. Educational Philosophy and Theory*, 54(6), 717-760. <https://doi.org/10.1080/00131857.2020.1777655>
- Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), 1753-1760. <https://doi.org/10.1177/1049732315617444>
- Martin, L. (2020). Foundations for good practice: The student experience of online learning in Australian higher education during the COVID-19 pandemic. *Tertiary Education Quality and Standards Agency (TEQSA)*. <https://www.teqsa.gov.au/sites/default/files/student-experience-of-online-learning-in-australian-he-during-covid-19.pdf>
- McDermott, D. (2012). Can we educate out racism? Medical

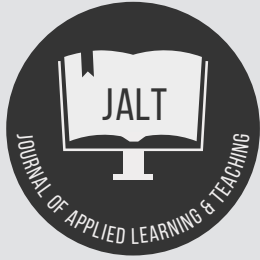
- Journal of Australia*, 197(1), 1-15. <https://doi.org/10.5694/mja12.10936>
- McGill, M., Turrietta, C., & Lal, A. (2021). Teaching health science students during COVID-19: Cross-hemisphere reflections. *Journal of University Teaching & Learning Practice*, 18(5). <https://doi.org/10.53761/1.18.5.3>
- McLeod, K., Thakchoe, S., Hunter, M. A., Vincent, K., Baltra-Ulloa, A. J., & MacDonald, A. (2020). Principles for a pedagogy of unlearning. *Reflective Practice*, 21(2), 183-197. <https://doi.org/10.1080/14623943.2020.1730782>
- McLeod, K., Moore, R., Robinson, D., Ozkul, D., Ciftci, S., & Vincent, K. (2020). Using the pluriverse concept to critique Eurocentrism in education. *Journal of Applied Learning & Teaching*, 3(S1), 30-33. <https://doi.org/10.37074/jalt.2020.3.s1.4>
- Molina-Azorin, J. F. (2016). Mixed methods research: An opportunity to improve our studies and our research skills. *European Journal of Management and Business Economics*, 25, 37-38. <http://dx.doi.org/10.1016/j.redeen.2016.05.001>
- Montelongo, R., & Eaton, P. W. (2019). Strategies and reflections on teaching diversity in digital space(s). In L. Kyei-Blankson, J. Blankson, & E. Ntuli (Eds.), *Care and culturally responsive pedagogy in online settings* (pp. 41-62). IGI Global.
- Moore, R., Rudling, E., Kunda, M., & Robin, S. (2021). Supporting casual teaching staff in the Australian neoliberal university: A collaborative approach. *Journal of Applied Learning & Teaching*, 4(2), 54-67. <https://doi.org/10.37074/jalt.2021.4.2.8>
- Morrison, A., Rigney, L.-I., Hattam, R., & Diplock, A. (2019). *Toward an Australian culturally responsive pedagogy: A narrative review of the literature*. Australia University of South Australia. <https://doi.org/10.13140/RG.2.2.13031.52641>
- Morse, J. M. (2000). Determining sample size. *Qualitative Health Research*, 10, 3-5. <https://doi.org/10.1177/104973200129118183>
- Mshigeni, S., Sarwar, E., & Leeunai, E. (2021). College students educational experiences amid COVID-19 pandemic. *Journal of Applied Learning & Teaching*, 4(1), 38-48. <https://doi.org/10.37074/jalt.2021.4.1.20>
- Mulrooney, H. M., & Kelly, A. F. (2020). COVID-19 and the move to online teaching: Impact on perceptions of belonging in staff and students in a UK widening participation university. *Journal of Applied Learning & Teaching*, 3(2), 17-30. <https://doi.org/10.37074/jalt.2020.3.2.15>
- Nyanjom, J., & Naylor, D. (2021). Performing emotional labour while teaching online. *Educational Research*, 63(2), 147-163. <https://doi.org/10.1080/00131881.2020.1836989>
- Ohito, E., & Oyler, O. (2017). Feeling our way toward inclusive counter-hegemonic pedagogies in teacher education. In L. Florian & N. Pantić (Eds.), *Teacher education for the changing demographics of schooling* (Vol. 2, pp. 183-198). Springer.
- Pawlowski, L. (2018) Creating a brave space classroom through writing. In S. Brookfield (Ed.), *Teaching race: How to help students unmask and challenge racism* (pp. 63-86). Jossey-Bass.
- Phan, P., Vugia, H., Wright, P., Woods, D. R., Chu, M., & Jones, T. (2009). Teaching note: A social work program's experience in teaching about race in the curriculum. *Journal of Social Work Education*, 45(2), 325-333. <https://doi.org/10.5175/JSWE.2009.200700056>
- Powell, Z., Davis, C., Yadav, R., & Gates, T. (2021). Using zoom in teaching human behavior during the COVID-19 pandemic: Experiences of Australian students. *Journal of Human Behavior in the Social Environment*, 32(8), 1101-1110. <https://doi.org/10.1080/10911359.2021.2002752>
- Saldaña, J. (2014). Coding and analysis strategies. In P. Leavy (Ed.), *The Oxford handbook of qualitative research* (pp. 581-605). Oxford University Press.
- Shevalier, R., & McKenzie, B. A. (2012). Culturally responsive teaching as an ethics-and care-based approach to urban education. *Urban Education*, 47(6), 1086-1105. <https://doi.org/10.1177/0042085912441483>
- Sjorberg, D., & McDermott, D. (2016). The deconstruction exercise: An assessment tool for enhancing critical thinking in cultural safety education. *International Journal of Critical Indigenous Studies*, 9(1), 28-48. <https://doi.org/10.5204/ijcis.v9i1.143>
- Stafford, V. (2020). EdTech review: Teaching through Zoom—what we've learned as new online educators. *Journal of Applied Learning & Teaching*, 3(2), 150-153. <https://doi.org/10.37074/jalt.2020.3.2.14>
- Sumer, M., Douglas, T., & Sim, K. (2021). Academic development through a pandemic crisis: Lessons learnt from three cases incorporating technical, pedagogical and social support. *Journal of University Teaching & Learning Practice*, 18(5). <https://doi.org/10.53761/1.18.5.1>
- Tang, H. H. H., & Tsui, C. P. G. (2018). Democratizing higher education through internationalization: The case of HKU SPACE. *Asian Education and Development Studies*, 7(1), 26-41. <https://doi.org/10.1108/AEDS-12-2016-0095>
- Terrell, S. (2012). Mixed-methods research methodologies. *The Qualitative Report*, 17(1), 254-280. <https://nsuworks.nova.edu/tqr/vol17/iss1/14>
- Thathsara, D., Maddumpatabandi, T. D., & Gamage, K. A. A. (2020). Novel coronavirus (COVID-19) pandemic: Common

challenges and responses from higher education providers. *Journal of Applied Learning & Teaching*, 3(2), 40-50. <https://doi.org/10.37074/jalt.2020.3.2.20>

Williams, R. (1999). Cultural safety—what does it mean for our work practice? *Australian and New Zealand Journal of Public Health*, 23(2), 213–214. <https://doi.org/10.1111/j.1467-842X.1999.tb01240>

Woodley, X., Hernandez, C., Parra, J., & Negash, B. (2017). Celebrating difference: Best practices in culturally responsive teaching online. *TechTrends*, 61, 470–478. <https://doi.org/10.1007/s11528-017-0207-z>

Copyright: © 2024. Robyn Moore, Kim McLeod and Natalia Maystorovich Chulio. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## A conceptual, strategic and implementation framework for the Scholarship of Learning and Teaching

---

Melanie Williams<sup>A</sup>

A

*Associate Dean (Scholarship), William Angliss Institute, Melbourne, Australia*

---

### Keywords

---

Reflective practice;  
scholarly practice;  
scholarship;  
Scholarship of Learning and Teaching (SoLT).

### Abstract

---

This paper outlines the conceptual, strategic and implementation framework that underpins the Scholarship of Learning and Teaching (SoLT) in an Australian non-university higher education provider. The standards for scholarly practice that lie at the heart of this framework are outlined, and the linkages between the related concepts of scholarship, research, professional development, and quality improvement are explored. An organisation-wide SoLT Strategy drives scholarly activities, while an implementation framework provides support and incentives that include guidelines and templates, mentoring and professional development, blanket ethics approval for the use of students' and staff's natural data with an online mechanism for managing consent, small funding grants, and awards for scholarly practice. SoLT is granted equal status with discipline-based research in promotion and tenure. The faculty produce and disseminate annual scholarly outputs and report annually on the impact of scholarly activities, which are monitored through academic governance committees. The initiative has proved successful, with 89 and 91 percent, respectively, of academics producing scholarly outputs that met the institute's standards for scholarly practice in the first two years of operation. The framework offers a comprehensive and coherent approach that may assist other higher education providers seeking to elevate the status and activity level of SoLT.

---

### Correspondence

[melanie.williams@angliss.edu.au](mailto:melanie.williams@angliss.edu.au) <sup>A</sup>

---

### Article Info

Received 19 December 2023  
Received in revised form 11 March 2024  
Accepted 11 March 2024  
Available online 4 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.29>

## Introduction

The Australian higher education regulatory framework, the *Higher Education Standards Framework (Threshold Standards) 2021*, known as the HES Framework, requires that all higher education providers engage with scholarship at both institutional and individual levels (Department of Education, Skills and Employment, 2021). Scholarship must contribute to course design and delivery and, where applicable, research and research training. Academics delivering higher education qualifications must engage in scholarship that informs their teaching and learning, and institutions are required to take a systematic approach to scholarship, including encouragement and support (TEQSA, 2022b).

Despite these regulatory requirements, many universities struggle to persuade academic staff to engage systematically in the Scholarship of Learning and Teaching (SoLT), not least because funding, promotion and tenure tend to be attached to discipline-based research rather than SoLT. On the other hand, for non-research-intensive higher education providers, engagement in SoLT can serve as a focus for building teaching quality. This paper sets out an institutional framework for scholarly practice at William Angliss Institute (WAI), an Australian mixed-sector specialist provider of higher education and vocational training in foods, hospitality, tourism, and events. Wheelahan et al. (2009) define mixed-sector providers as those in which less than 20% of provision is in one sector, as distinct from dual-sector providers, which have a more even spread. In its first two years of implementation, the framework has seen an extraordinary degree of success in enabling initiatives designed to enhance the student learning experience and outcomes on the one hand, and in producing and disseminating quality scholarly outputs on the other.

## The conceptual framework

### Definitions

The current guidance note on scholarship produced by Australia's higher education regulatory authority, the Tertiary Education Quality and Standards Agency (TEQSA), defines scholarship as referring to "those activities concerned with gaining new or improved understanding, or appreciation and insights into a field of knowledge, or engaging with and keeping up to date with advances in the field" (TEQSA, 2022b, p. 1). It shies away from identifying any particular model of scholarship, stating that the Tertiary Education Quality Standards Authority (TEQSA), the Australian higher education regulator, "recognises there is no singular definition of scholarship and acknowledges that providers may utilise various approaches to organise the full range of their scholarly activities" (TEQSA, 2022b, p. 1).

However, WAI's Framework for Scholarly Practice was influenced by TEQSA's earlier (2018) guidance note on scholarship, which referenced Boyer's (1990) depiction of four discrete, yet interdependent forms of scholarship: discovery (pure research), integration (bringing cross-disciplinary insights to bear or translating specialist scholarship for lay

audiences), application (where every-day societal problems set the research agenda – conceptually superseded by the notion of 'engagement' (Boyer, 1996)) and teaching.

Boyer's (1990) work was seminal. Hitherto, promotion and tenure in universities had been based solely on research output. However, Boyer pointed out that academics perform a broader range of functions than just pure research (especially in institutions such as the 'land grant' universities in the United States, which had been established to fulfil a different mission) and that these other forms of academic work were equally deserving of recognition. While most research-intensive universities continue to value pure research over Boyer's other forms of scholarship, his framework has been widely adopted in other types of higher education institutions. For teaching-intensive institutions such as WAI, Boyer's framework is particularly noteworthy because it elevated the status of teaching to a form of scholarship for the first time.

The scholarship of teaching has come to include learning, noting the shift from a teaching to a learning paradigm (Barr & Tagg, 1995). This focus on learning is emphasised by WAI's preference for putting learning first in the term 'Scholarship of Learning and Teaching' (SoLT), rather than the more common 'Scholarship of Teaching and Learning' (SoTL). Whereas TEQSA's 2018 guidance note referred to the scholarship of teaching as that which 'promotes active and critical learning in students based on advances in a discipline or in knowledge about effective teaching and learning and course design practices in a field' (TEQSA, 2018, p. 2), the 2022 update does not differentiate or define the scholarship of teaching.

While acknowledging that under Boyer's framework, 'scholarship' is the parent concept that encompasses the four different forms, at WAI, the notion of 'scholarship' pertains specifically to SoLT. This is primarily because discovery, engagement and integration are accounted for separately in the institute's Enterprise Bargaining Agreement (EBA) under the banner of 'research' for the purposes of workload allocation. WAI's Framework for Scholarly Practice defines scholarship as 'the practice of taking a planned, rigorous, and reflective approach to investigating an aspect of practice and using the results to inform and improve learning and teaching' (William Angliss Institute, 2017, p. 1). In an institution closely connected with industry, 'practice' may refer to industry or professional practice, discipline-based research practice or teaching practice. The important point is that whatever practice forms the focus of investigation, the findings are used to improve students' active and critical learning.

Hence, WAI's approach to scholarship is consistent with Tight's (2018) assessment of an attempt at "recognising the importance of taking a critical and research-based approach to teaching and learning, and, in doing so, attempts to elevate the status of the teaching role in comparison to research" (p. 2).



## **The relationship between professional development and SoLT**

WAI's definition of scholarship contrasts sharply to Boshier's (2009) and Coderch's (2023) assertions that the SoTL literature frequently conflates SoTL with other activities, with the three most common of these being attending conferences, workshops and seminars; taking training courses; and doing peer review. None of these three activities would meet Shulman's (2000) widely adopted criteria for rigorous scholarship – that is to say, being made public, available for peer review and critique according to accepted standards, and able to be reproduced and built on by other scholars. However, Boshier (2009) condemns what he terms as "uncritical over-reliance on peer review as the mechanism for measuring scholarship" (p. 1). He goes on to claim that Boyer's definition was conceptually flawed, in that it never clearly articulated what distinguished the scholarship of teaching from teaching excellence.

TEQSA's (2018) guidance note explicitly associates scholarship with advances in knowledge and practice. That is, knowledge and practices that are new to the field and/or its pedagogy, not just new to the individual. As such, attending conferences, workshops and seminars, and taking training courses constitute professional development (PD), not scholarship. Thus, regardless of whether there may be an over-reliance on peer review as a measure of quality, a defining difference between scholarship and PD is that scholarship necessarily involves the public dissemination of new knowledge (Glassick et al., 1997; Shulman, 2000; Williams et al., 2013).

Nonetheless, WAI does acknowledge there can be a relationship between these activities and SoLT in particular circumstances. PD activities that are extended through 1) implementation in the classroom, 2) evaluation of their impact on active and critical learning, and 3) public dissemination of the results would fulfil WAI's scholarship requirements.

## **The relationship between quality improvement/assurance of learning and SoLT**

In a similar way, activities normally associated with routine quality improvement (QI), such as student subject evaluation, subject and course review, or what Hall and Ko (2006) define as "the process by which student learning outcomes are measured against specific course goals" (p. 1), may be used as a basis for scholarship if they are extended in the three ways outlined above.

## **The relationship between research and SoLT**

TEQSA's (2022a) guidance note on research and research training defines research as 'the creation of new knowledge and/or the use of existing knowledge in a new and creative way by a higher education provider so as to generate new concepts, methodologies, inventions and understandings' (p. 1). This definition could equally apply to scholarship, as defined at WAI. The academic literature also acknowledges

an overlap between research and SoLT. For instance, Ling (2020) concludes that:

an academic activity may involve both scholarship of teaching and learning and education research, provided that, *inter alia*, it involves systematically investigating a contentious issue or a gap in current understandings of education, in a form sufficient to warrant conclusions that have the potential to contribute to current understandings of pedagogy or other aspects of education (p. 56).

Canning and Masika (2022) go further, calling for the complete abandonment of SoLT in favour of asserting the value of higher education research. However, Healy et al. (2020) draw a useful distinction between educational research and SoLT. For these authors, whereas the primary goal of educational research is to generate generalisable knowledge, the fundamental purpose of SoLT is to improve teaching and learning for the group of students being studied.

Such debates centre on the relationship of SoLT to educational research. These issues are avoided at WAI through making a clear distinction between research, which pertains to WAI's specialist disciplinary domains, and scholarship, pertaining to learning and teaching. Hence, research at WAI is defined as "the generation of new knowledge through original investigation that leads to advances in the disciplinary knowledge and professional practices associated with the domains of foods, hospitality, tourism and events" (William Angliss Institute, 2022, p. 14).

As stated above, for the purposes of workload allocation under the EBA, research can incorporate Boyer's scholarships of engagement and integration related to the domains mentioned in the earlier paragraph. The workload percentage allocated to research is negotiable, but typically defaults to 20 per cent. However, just as with QI activities, disciplinary research may be used as a basis for SoLT if it is extended as outlined above. That is, the disciplinary research in and of itself would not count as scholarship but if the outcomes were applied in the classroom, their impact on promoting active and critical learning were evaluated and the results of this evaluation publicly disseminated, this would count as scholarship.

The linkages between the four interrelated activities of professional development, quality improvement, research and SoLT are shown in Figure 1. As explained, any activities in the diagram may contribute to scholarship only in so far as they contribute to promoting active and critical learning based on advances in discipline, professional practice or knowledge about learning and teaching and course design – and ultimately, lead to scholarly outputs. This relationship is represented by arrows and/or overlaps with the SoLT circle.

## **The strategic framework**

WAI develops successive institute-wide, three-year strategies for SoLT, which sit under a broader strategy for educational excellence, which is in turn subordinate to the Institute's

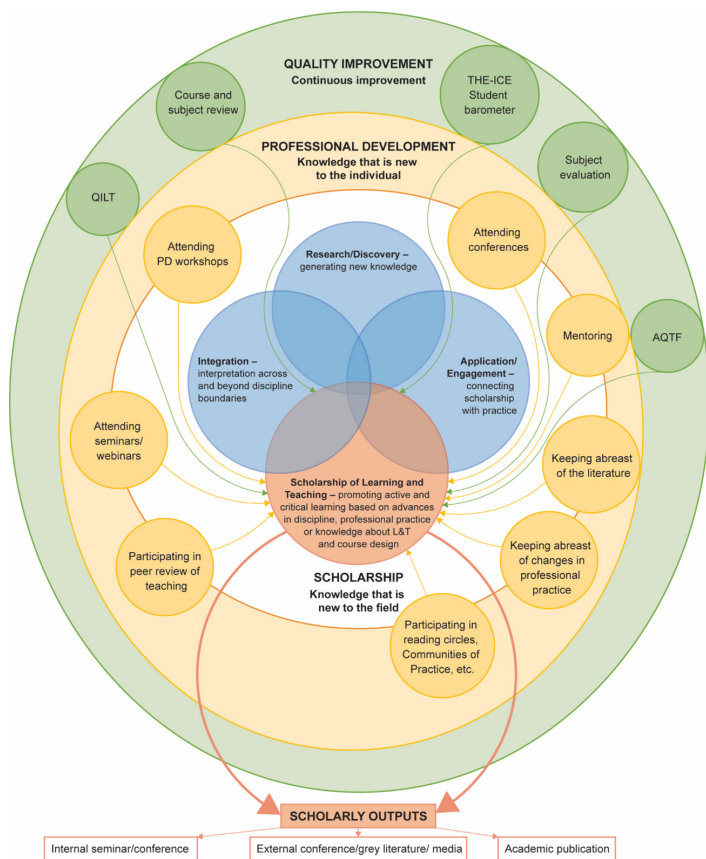


Figure 1: The relationships between QI, PD, research and scholarship.

strategic plan. The goals of the current SoLT Strategy have been to consolidate the work that began in 2017 of embedding scholarly practice as a systemic approach to improving quality and driving innovation in learning and teaching and to strengthen the monitoring of its quality and impact through academic governance structures and processes. The goals are broken down into six objectives:

- (1) To advance evidence-based approaches to improving the quality of learning and teaching.
- (2) To advance the design, development, deployment and evaluation of innovations in curriculum and pedagogy.
- (3) To evaluate the impact of the deployment of pedagogies consistent with WAI's constructivist educational philosophy.
- (4) To promote WAI's thought leadership through disseminating the outcomes of innovation and scholarly practice.
- (5) To build the capacity of the WAI teaching workforce in SoLT.
- (6) To enhance and maintain governance oversight of the quality and impact of scholarship.

Progress made towards these objectives is reported quarterly through the Research and Scholarship Committee, a standing committee of the highest academic governance

committee. SoLT Strategy progress is reported to this committee biannually.

## The implementation framework

The SoLT Strategy is operationalised through an implementation framework. The single most crucial factor in ensuring engagement in scholarship at WAI is the inclusion of a mandatory minimum five per cent workload allocation for scholarship for all academic staff, which is enshrined in the Enterprise Bargaining Agreement (EBA). Academics who wish to focus on teaching may reduce discipline-based research for SoLT and negotiate for up to an additional 10% of their workload for scholarship. Prior to the mandatory workload allocation in the current EBA, engagement in scholarship tended to be ad hoc, undocumented, undertaken by few staff and did not fulfil all the WAI Standards for Scholarly Practice. In contrast, in the first year of implementation of the current framework, 89% of academics produced and disseminated scholarly outputs that met the standards, with that figure rising to 91% in the second year.

## Standards for scholarly practice

At the heart of the framework is a set of standards adapted from Glassick et al. (1997) and Williams et al. (2013), which act as a lens through which the other elements of the framework are mediated. They provide the benchmark against which all elements are measured. The standards were originally developed through Glassick et al.'s (1997) analysis of documentation from American universities, such as guidelines for hiring, promotion and tenure, criteria used by academic publishers and grant agencies to evaluate submissions, and so on. However, in the research of SoLT in a range of higher education contexts in Australia, Williams et al. (2013) concluded that certain assumptions that pertain to research and scholarship in universities do not necessarily hold in mixed-sector institutions such as WAI, in which a cultural legacy from vocational education and training (VET) tends to prevail. These authors argued that the standards developed by Glassick and his colleagues should be augmented for mixed-sector institutions. For example, it should be made explicit that the best practice in scholarship involves additional factors, such as collaboration, critical analysis and synthesis, theory-informed practice and making work public. These factors cannot be assumed in institutions built on a legacy of VET custom and practice.

Accordingly, the WAI standards for scholarly practice were adopted with these insights in mind. The standards are laid out in Table 1.

## Planning

Each teaching academic is required to develop and implement a three-year scholarship plan. All plans are checked by the Associate Dean (Scholarship) for alignment with the Standards for Scholarly Practice before being submitted to the Research and Scholarship Committee for approval.

Table 1: Standards for scholarly practice adapted from Glassick et al. (1997) and Williams et al. (2013).

Features of scholarly work	In your investigation, do you...
<b>Clear goals</b> – at the start, outlining what you plan to achieve	<ul style="list-style-type: none"> <li>➤ Explain the project clearly and identify important questions in the field?</li> <li>➤ Define a clear purpose and achievable objectives for your work?</li> </ul>
<b>Adequate preparation</b> – relating your work to what is already known about the topic and gathering the resources needed for the job	<ul style="list-style-type: none"> <li>➤ Locate your work in the context of current and emerging industry/ disciplinary/ pedagogical knowledge/ practice/ research in your field?</li> <li>➤ Bring together the resources, information and skills necessary to move the work forward?</li> </ul>
<b>Appropriate methods</b> – choosing and applying the best way to achieve your goals and keeping records of progress	<ul style="list-style-type: none"> <li>➤ Use a systematic and planned approach that is appropriate to the goals?</li> <li>➤ Apply the methods selected in a rigorous and ethical manner that is responsive to changing circumstances?</li> <li>➤ Maintain records of process and outcomes?</li> </ul>
<b>Collaboration</b> – working and learning with others	<ul style="list-style-type: none"> <li>➤ Engage with a range of stakeholders?</li> <li>➤ Draw on specialist expertise and advice?</li> <li>➤ Engage in shared reflection?</li> </ul>
<b>Critical analysis and synthesis</b> – questioning what is going on and bringing things together to make sense of them	<ul style="list-style-type: none"> <li>➤ Challenge existing knowledge, assumptions and ideas?</li> <li>➤ Bring together your findings to draw conclusions within a theoretical framework?</li> <li>➤ Support your claims with evidence and sound argument?</li> </ul>
<b>Significant results</b> – making a difference	<ul style="list-style-type: none"> <li>➤ Contribute new knowledge, the new application of knowledge or improved practice in the field?</li> <li>➤ Offer students the opportunity for innovative engagement with their future profession?</li> <li>➤ Open up additional areas for further exploration?</li> </ul>
<b>Making knowledge public</b> – sharing new knowledge with others so that it can be critiqued, built upon and improved	<ul style="list-style-type: none"> <li>➤ Open your practice to peer review and stakeholder feedback?</li> <li>➤ Communicate your message clearly through teaching, presentation, publication or exhibition?</li> </ul>
<b>Reflective critique</b> – reflecting on the strengths, weaknesses and limits of your work in order to do better next time	<ul style="list-style-type: none"> <li>➤ Identify the influences and assumptions that you bring to the work?</li> <li>➤ Reflect on both the processes and outcomes?</li> <li>➤ Bring a breadth of evidence to the review of your work?</li> <li>➤ Use critical evaluation to improve the quality of future work?</li> </ul>

Since a five-per-cent workload equates to less than 12 days per year, it is suggested that the plan consists of just one project over the three years to enable a substantial undertaking of sufficient depth. Typically, the first year is spent doing a literature review, professional development activities related to the topic of investigation, and designing a teaching intervention that is informed by these activities. The second year typically involves implementing the intervention and gathering data to evaluate its effectiveness. The third year may be spent writing up the project as a journal article or other publication and preparing the next three-year plan. Since there is an expectation that the outcomes of each year's scholarly activity will be disseminated, much of the analysis and writing has already been done in the previous two years before writing up the final article.

Academics are asked to anticipate what impact their scholarship may have (more on this below) and to identify the type of evidence could support this in their plan, at least in a preliminary way. In the past, academics have typically relied primarily on student evaluations as evidence of impact, despite their highly contested value (Bartkowiak-Theron et al., 2020). Designing impact evaluation into the planning of the project reminds academics to consider more broadly what might constitute appropriate evidence and to gather it along the way, thus providing more reliable measures and making annual reporting on impact easier.

## Indicative developmental hierarchies of scholarly activities and outputs

The three-year planning cycle facilitates a developmental approach to scholarship. This is reflected in increasing rigour of activities normally engaged in year on year, and the annual scholarly outputs that staff are expected to produce. Generally, the first-year literature review is presented to internal staff at an annual Scholarship Symposium. In addition to presenting internally in the second year, those academics who wish to and are financially supported to present a working paper at an external conference or they may produce a brief publication for the grey literature. Given WAI's close ties with industry and its aspirations as a thought leader for the hospitality and tourism industry, industry publications are highly valued. Scholarly outputs typically culminate in an academic publication in the third year.

Academics may deviate from this indicative pathway: it is not unusual, for example, for them to present their first year's work as a working paper at an external conference or to produce some other output. Nonetheless, Table 2 below sets out an indicative developmental hierarchy of scholarly activities and outputs.

Table 2: Indicative levels of activity and associated outputs by increasing rigour.

	Indicative activity types	Typical outputs
Level 1	<ul style="list-style-type: none"> <li>➤ Undertake PD activities to learn about advances in knowledge of the discipline, professional practice, learning &amp; teaching or course design</li> <li>➤ Peer review of teaching</li> <li>➤ Reflective practice individually and/or with others</li> <li>➤ Pilot &amp; get feedback on new pedagogical approach</li> <li>➤ Develop presentation for internal audience</li> </ul>	Internal presentation
Level 2	<ul style="list-style-type: none"> <li>➤ Design strategy to evaluate impact of new pedagogical approach on active and critical learning</li> <li>➤ Implement new pedagogical approach</li> <li>➤ Gather &amp; analyse data about the impact on active and critical learning (evaluation)</li> <li>➤ Develop presentation for external audience</li> <li>➤ Write for grey publication</li> </ul>	External presentation/ grey publication/ media learning
Level 3	<ul style="list-style-type: none"> <li>➤ Write for academic publication</li> </ul>	Academic publication (Book chapter or minimum B ranked journal)

In addition, there is a hierarchy of outputs according to academic classification, generally with encouragement for senior staff to aim for publication in more highly ranked journals. However, it should be stressed that these guidelines are indicative only: WAI's close ties to industry and aspiration to be a thought leader in this space imply that impacts on industry custom and practice are highly valued and may take precedence over more traditional academic outputs.

## Monitoring and reporting impact and outcomes

Progress on scholarship plans is monitored annually by the Research and Scholarship Committee. In an adaptation of Hinton's (2016) *Impact Management Planning and Evaluation Ladder (IMPEL)*, which was originally developed to measure research impact, academics are asked to

report on the impact of their scholarship in progressively wider spheres of influence. These include changes in the project team's thinking and understanding, changes to their teaching practice, benefits to the students' learning experience, and contributions to the field. This is similar in scope to Simmons's later (2020) 4M Framework that was developed specifically for measuring the impact of SoLT: micro (individual/researcher) level, meso (departmental) level, macro (institutional) level and mega (discipline/national) level.

Reporting on impact is accompanied by evidence to support these claims. As mentioned, the pre-identification of potential evidence of impact at the planning stage greatly assists when it comes to reporting actual impact. Where possible, a significant proportion of the evidence is generated through analysis of students' natural data in order to minimise the impost on students.

Reflective practice is highly valued as part of scholarly culture at WAI, with reflective critique being the most important feature in the institute's standards for scholarly practice (Glassick et al., 1997). Academics are therefore asked to reflect critically on their scholarship, not only to identify improvements they could make to their scholarly projects, but also to track their own learning and development in SoLT. If required, the three-year plans can be adjusted in light of these reflections and changes logged with the Research and Scholarship Committee.

### **Institutional support**

Under the HES Framework, higher education providers must provide encouragement and support for scholarship at an institutional level (Department of Education, Skills and Employment, 2021). At WAI, this support takes multiple forms, in addition to the workload allocation for scholarship outlined above. Such support includes mentoring, guidelines and templates, provision of formal and informal professional development, blanket ethics approval with an online mechanism for managing consent for the use of natural student and staff data, provision of small grant funds seminars, webinars and an annual scholarship symposium, awards for scholarly practice, and equal recognition of scholarship with disciplinary research for the purposes of promotion and tenure.

### **Mentoring, professional development and the dissemination of findings**

WAI employs a dedicated full-time Level D academic to lead, mentor and support staff to engage in scholarship across the institution, including in vocational education where there is no regulatory requirement to undertake scholarship. In addition to the oversight of all SoLT-related scholarly activity, the Associate Dean (Scholarship) provides professional development in SoLT and guidance and mentoring in the development and implementation of the three-year scholarship plans. This on-demand assistance that includes providing guidance and feedback on draft outputs, on a one-on-one or small team basis, is available

throughout the life of the projects. Where it becomes evident that there is a collective need for a particular aspect of scholarship, a professional development webinar may be scheduled for all interested staff.

There is provision for external experts on SoLT to offer seminars as part of WAI's annual Research and Scholarship Seminar Program. In addition, WAI schedules an online symposium in December each year to provide a forum for staff to disseminate the findings from the year's scholarship activities. The symposium is open to all Institute staff, both as presenters and audience. The unpublished outputs (or as a minimum, an abstract thereof) are housed on the organisation's Intranet so that other staff may access the learnings to inform their own teaching practice.

### **Management of ethics for the use of student and staff data**

Following the University of Tasmania's Curriculum Evaluation Framework (Kelder & Carr, 2017; Kelder et al., 2017), WAI's Research Ethics Committee has granted blanket ethics approval for the use of student and staff data for SoLT purposes. Ethics approval only pertains to the use of 'natural' data that has been produced in the normal course of undertaking a course, such as assessments and online or classroom activities that have been uploaded into the Learning Management System (LMS). The purpose of this is to use as much existing data as possible to minimise the workload imposed on students when staff undertake their SoLT activities. If academics need to generate additional data through surveys, focus groups or interviews with students, this requires a separate, full ethics application through the Research Ethics Committee as the norm.

Consent is managed online via a portal in the LMS. The usual ethics information, such as a plain language statement that is required for informed consent, is made available in this portal. Students only need to give permission once to cover their whole course. However, they can update their consent status at any time, in real time, up to two weeks after results are finalised for each semester. After that date, the data is anonymised by an independent data manager and made available to academics. Individual students can be traced via their student number to enable longitudinal and comparative studies, but the data matching is handled by the data manager at the back end of the LMS and academics have no access to the identified data.

'Natural' staff data includes anything that is produced as part of delivering a course, such as feedback on assessment, session plans, scholarship reports and so on. While the blanket ethics approval also covers the use of such data, the academics can upload whichever documents they are willing to make available for research purposes into the online portal. As a result, not all relevant documents are housed and readily accessible in the LMS.

## Grants and awards

Micro grants of up to \$100 per person, per project, per year are available to support minor expenses, such as interview transcription, catering or software licenses. Up to three years' worth of grant allocation may be taken at a time.

Applications for Awards for Scholarly Practice are opened in November each year. Applicants are invited to submit a scholarly output, such as a journal article or conference paper, accompanied by an exegesis explaining how the artefact meets WAI's Standards for Scholarly Practice. A template based on the standards is provided, along with the scoring rubric that is based on these same standards used to assess the applications.

There are three categories of award: Award for Merit in Scholarly Practice, Award for Excellence in Scholarly Practice and Scholar of the Year Award. To receive any award, all standards must be met. Applications are scored on a five-point scale for each criterion, with one denoting 'just achieved' and five denoting 'outstanding'. An average equal to or less than two per criterion achieves 'Merit'. 'Excellence' is awarded to applications scoring an average greater than two per criterion. Scholar of the Year is awarded to the applicant who receives the highest score in the excellence category.

A panel, comprising the Associate Dean (Scholarship) and two other members of the Research and Scholarship Committee, assesses the applications. Prizes consist of funds towards professional development, including conference attendance. The Scholar of the Year receives funds for excellence and the Scholar of the Year, receives effectively double the prize money.

## Promotion

In contrast to many universities where research outputs are privileged over SoLT, there is no such distinction when it comes to promotion and tenure at WAI: scholarship and research outputs are equally valued. However, a hierarchy applies to the 'quality' of publications from highly ranked academic journals through to grey literature when it comes to expected outputs at the various employment classification levels. These priorities were entrenched in the EBA at a time when WAI aspired to become a University of Specialisation (Williams, 2018). This provider category no longer exists in Australia and WAI's priorities have shifted accordingly. In the next round of EBA negotiations, it is possible that the balance may move towards a higher value for outputs that demonstrate thought leadership for industry, reflecting WAI's shift in strategic emphasis.

## Governance and quality oversight

Scholarly activity at both institutional and individual levels is overseen by the Research and Scholarship Committee, a standing committee of the equivalent of an Academic Board. The institute-wide three-year SoLT Strategy is approved by the Research and Scholarship Committee, and progress

towards meeting its objectives is reported quarterly.

Academics' individual three-year scholarship plans are approved by this same committee and their impact is reported to it annually. Annual scholarly outputs and impact reports are checked by the Associate Dean (Scholarship) for alignment to the WAI Standards for Scholarly Practice. Feedback and developmental support are offered where the standards have not been met. Failure to meet the annual requirements is referred to the academic's line manager for performance management. The impact of academics' scholarship and the production and dissemination of scholarly outputs are reported to the institute's highest academic governance body annually. Figure 2 below shows the system of frameworks governing and supporting SoLT at WAI.



Figure 2: Summary of WAI's institutional frameworks governing and supporting SoLT.

## Future improvements

While the current arrangements have resulted in a high level of engagement in scholarship, quality scholarly outputs and measurable impact – as evidenced by the percentage of academic staff who have produced scholarly outputs that meet WAI's Standards for Scholarly Practice and by the empirical data supporting claims to impact, there is room for improvement.

To date, academics have been free to select their own area of focus for their scholarship. As an initial position, engaging in an area of personal interest has enhanced motivation while less experienced academics gain understanding and proficiency in SoLT. However, this may need to change as WAI's scholarship matures. It has inhibited WAI's ability to use scholarship to advance strategic initiatives because the available workload allocations have been taken up with individual and small team projects. For instance, a recent overarching learning and teaching strategy, *The Strategic Framework for Educational Excellence* (William Angliss Institute, 2022) was introduced at the commencement of the latest three-year SoLT planning cycle. In it, the Institute has committed to constructivist educational philosophy and specific pedagogical approaches within that philosophy. Engaging academics in a broader, collective initiative to build on constructivist pedagogies as a focus for SoLT, as compared to choosing their own area of focus, would advance the Institute's strategic goals, while simultaneously enhancing students' active and critical learning.

Furthermore, feedback from faculty has consistently indicated that the expected outputs for a five percent workload are too high to be achieved within the allotted time allocation. WAI would be reluctant to retreat from the

significant outcomes achieved through its SoLT initiative, so a higher workload allocation may need to be negotiated in the next round of enterprise bargaining.

## Conclusion

The WAI Framework for Scholarly Practice provides a comprehensive and integrated array of enabling mechanisms, guidance, support, recognition, reward and governance oversight to ensure that the Institute not only meets its regulatory obligations with regard to scholarship, but also engenders a culture in which making improvements to students' active and critical learning is central, and SoLT is valued and celebrated. The framework offers ideas that can be taken up in other higher education settings where there is a desire to elevate the status and activity dedicated to SoLT. The elements of the framework may be adopted individually or as part of a comprehensive approach. They may be of particular value in supporting greater engagement in scholarship in university settings, where SoLT often comes a distant third behind disciplinary research and teaching. Indeed, the framework shows other institutions how to enhance the impact and visibility of SoLT, providing a rigorous pathway to recognition for teaching-focused academics and demonstrating how to produce a measurably positive impact on the active and critical learning of students.

## References

- Barr, R., & Tagg, J. (1995). From teaching to learning - a new paradigm for undergraduate education. *Change*, 27(6), 12-25. <https://doi.org/10.1080/00091383.1995.10544672>
- Bartkowiak-Theron, I., McShane, A. L. J., & Knight, M. G. (2020). Departing from anonymous and quantitative student feedback: Fostering learning and teaching development through student evaluations. *Journal of Applied Learning and Teaching*, 3(S1), 118-128. <https://doi.org/10.37074/jalt.2020.3.s1.16>
- Boshier, R. (2009). Why is the scholarship of teaching and learning such a hard sell? *Higher Education Research & Development*, 28(1), 1-15. <https://doi.org/10.1080/07294360802444321>
- Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Jossey-Bass. <https://www.umces.edu/sites/default/files/al/pdfs/BoyerScholarshipReconsidered.pdf>
- Boyer, E. L. (1996). The scholarship of engagement. *Bulletin of the American Academy of Arts and Sciences*, 49(7). <https://doi.org/10.2307/3824459>
- Canning, J., & Masika, R. (2022). The scholarship of teaching and learning (SoTL): The thorn in the flesh of educational research. *Studies in Higher Education*, 47(6). <https://doi.org/10.1080/03075079.2020.1836485>
- Coderch, M. (2023). Beyond the 'research vs. scholarship' dichotomy: The emergence of a new category of academic staff. *International Journal of the Scholarship of Teaching and Learning*, 17(2, Article 4), 1-8. <https://doi.org/10.20429/ijstol.2023.17204>
- Department of Education, Skills and Employment. (2021). *Higher education standards framework (Threshold Standards) 2021*. <https://www.legislation.gov.au/Details/F2022C00105>
- Glassick, C. E., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship assessed: Evaluation of the professoriate*. Jossey-Bass.
- Hall, O., & Ko, K. (2006). Learning assurance using business simulations to executive management education. *Developments in Business Simulations and Experiential Learning*, 33.
- Healy, M., Matthews, K. M., & Cook-Sather, A. (2020). *Writing about learning and teaching in higher education*. Elon Center for Engaged Learning. [https://www.centerforengagedlearning.org/wp-content/uploads/2020/09/writing-about-learning\\_chapter-1.pdf](https://www.centerforengagedlearning.org/wp-content/uploads/2020/09/writing-about-learning_chapter-1.pdf)
- Hinton, T. (2016). *The Impact Management Planning and Evaluation Ladder (IMPEL)*. Australian Government Department of Education. <https://www.education.gov.au/learning-and-teaching/resources/impact-management-planning-and-evaluation-ladder-impel>
- Kelder, J.-A., & Carr, A. (2017). Embedding evaluation and scholarship into curriculum and teaching: The curriculum evaluation research framework. In A. Horsted, P. Bartholomew, J. Branch, & C. Nygaard (Eds.), *New innovations in teaching and learning*. Libri Publishing.
- Kelder, J.-A., Carr, A., & Walls, J. (2017). Curriculum evaluation and research framework: Facilitating a teaching team approach to curriculum quality. *40th HERDSA annual international conferenc*. Sydney. <https://hdl.handle.net/102.100.100/523354>
- Ling, P. (2020). The scholarship of teaching and learning and higher education research: Boundaries and implications. In L. Ling & P. Ling (Eds.), *Emerging methods and paradigms in scholarship and education research*. IGI Global. <http://dx.doi.org/10.4018/978-1-7998-1001-8.ch003>
- Shulman, L. S. (2000). *Fostering a scholarship of teaching and learning*. Annual Louise McBee Lecture, University of Georgia. [https://www.researchgate.net/publication/234715873\\_Fostering\\_a\\_Scholarship\\_of\\_Teaching\\_and\\_Learning](https://www.researchgate.net/publication/234715873_Fostering_a_Scholarship_of_Teaching_and_Learning)
- Simmons, N. (2020). The 4M framework as analytic lens for SoTL's impact: A study of seven scholars. *Teaching and Learning Inquiry*, 8(1), 76-90. <https://doi.org/10.20343/teachlearninqu.8.1.6>
- TEQSA. (2018). *Guidance note: Scholarship*. Australian Government. [http://www.teqsa.gov.au/sites/default/files/GuidanceNote\\_Scholarship2.3.pdf](http://www.teqsa.gov.au/sites/default/files/GuidanceNote_Scholarship2.3.pdf)
- TEQSA. (2022a). *Guidance note: Research and research training*. Australian Government. <https://www.teqsa.gov.au/guides-resources/resources/guidance-notes/guidance->

note-research-and-research-training

TEQSA. (2022b). *Guidance note: Scholarship*. Tertiary Education Quality and Standards Agency. <https://www.teqsa.gov.au/guides-resources/resources/guidance-notes/guidance-note-scholarship>

Tight, M. (2018). Tracking the scholarship of teaching and learning. *Policy Reviews in Higher Education*, 2(1), 61-78. <https://doi.org/10.1080/23322969.2017.1390690>

Wheelahan, L., Moodie, G., Billett, S., & Kelly, A. (2009). *Higher education in TAFE*. NCVET. [https://www.ncver.edu.au/\\_data/assets/word\\_doc/0015/6900/higher-education-in-tafe-2167.doc](https://www.ncver.edu.au/_data/assets/word_doc/0015/6900/higher-education-in-tafe-2167.doc)

William Angliss Institute. (2017). *Framework for scholarly practice*. William Angliss Institute. <https://portal.angliss.vic.edu.au/research/Scholarship%20Planning%20and%20Reporting/Framework%20for%20Scholarly%20Practice%202021.pdf>

William Angliss Institute. (2022). *The strategic framework for educational excellence*. William Angliss Institute.

Williams, M. (2018). Plausible futures: Transforming ourselves, transforming our industry. In S. Beeton & A. Morrison (Eds.), *The study of food, tourism, hospitality and events: 21st-Century approaches* (pp. 257-267). Springer Singapore. [https://doi.org/10.1007/978-981-13-0638-9\\_23](https://doi.org/10.1007/978-981-13-0638-9_23)

Williams, M., Goulding, F., & Seddon, T. (2013). *Towards a culture of scholarly practice in mixed-sector institutions*. NCVET. <https://www.ncver.edu.au/research-and-statistics/publications/all-publications/towards-a-culture-of-scholarly-practice-in-mixed-sector-institutions>

Copyright: © 2024. Melanie Williams. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Micro-credentials in higher and vocational education: An innovation or a disruption? A review of the literature

---

Jeremy Hanshaw<sup>A</sup>

A

*Transdisciplinary International Tertiary Education Specialist, Academic Sessional Advisor, University of Leicester*

---

### Keywords

Disruption;  
higher education;  
innovation;  
micro-credentials;  
vocational education.

---

### Abstract

This article examines the literature on micro-credentials within the context of higher education and vocational education. It considers whether they are an innovative force for good or a disruptive force for evil. Are they, as the literature suggests, a means of creating agency, affording equity, access, and participation in higher and vocational education for those who otherwise lacked the time, money, opportunity, or confidence to apply for further study or/and felt disenfranchised from the learning experience, or/and found the whole concept of a qualification daunting? Are they, as posited in the literature, an excellent conduit to higher and vocational education for those wishing to sample an academic or vocational subject without committing to a full degree course? Or are they, as pre-supposed in other literature, an over-simplistic alternative to the traditional academic credential, a cynical attempt to dumb down knowledge, turning higher and vocational education into a series of stackable credentials aimed at satisfying the job market, and the neoliberal thirst for more and more dollars to fund our institutions, but failing to meet the finer subtleties of the academic experience? I examine and critique the literature around this debate and argue how we might harness micro-credentials to sustain innovation and disruption positively, leveraging them to move forward within education in general and higher and vocational education in particular.

---

### Correspondence

[jeremy.hanshaw@gmail.com](mailto:jeremy.hanshaw@gmail.com)<sup>A</sup>

---

### Article Info

Received 29 January 2024

Received in revised form 1 March 2024

Accepted 18 April 2024

Available online 22 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.39>



## Introduction

I start by briefly defining micro-credentials and considering some of their affordances within the higher education (HE) and vocational education and training (VET) sectors. The second section addresses the disruptive element of micro-credentials, reviewing and critiquing the idea posited by some of the literature that they are negatively disruptive and a threat to the academy. The third section looks at the potential of micro-credentials to create agency and contribute to equity, access, and participation in higher education to be positively disruptive, as evidenced in the current literature. Finally, based on this debate in the literature, recommendations are made on how we might leverage micro-credentials in HE and VET in the future.

## Definitions

I will start by discussing the definitions of micro-credentials (MCs) and micro-credentialing (MCg). The literature on defining MCs is extensive yet increasingly, and rightly, narrow as we systematically agree upon a definition through refining and defining our terms, which previously bound educators and innovators in respect of MCg (Oliver & UNESCO, 2022).

As recently as five years ago, practitioners did not have an agreed definition of MCs. Educators, in our discussions, did not know what they were, did not have a clear distinction between an MC and a digital badge, did not know whether they should contain learning of subject matter as well as the earning of a qualification, or whether they could or should be stacked. Five years later, vocational and higher education sectors arguably agree on what MCs are but disagree on how they might best be deployed to benefit learners. In the following paragraphs, I shall consider the former and the latter.

I will now consider the numerous characteristics of an MC; these characteristics will lead us to numerous affordances of MCs. An MC is, as the name suggests, micro, a small unit of learning that is credentialed, i.e., assessed. The European Commission defines MCs as:

A documented statement awarded by a trusted body to signify that a learner upon assessment has achieved learning outcomes of a small volume of learning against given standards and in compliance with agreed quality assurance principles. (European Commission, 2020)

The Commission states that MCs demonstrate credit volume and are aligned with national and European Qualification frameworks. They may be offered face-to-face, online, or blended learning means and can be formal or informal. It belongs to the owner and is sharable, portable, stand-alone, or part of a more extensive portfolio, triggering an award or digital badge (European Commission, 2020). The New Zealand Qualifications Authority (2020) defines them as “a sub-set of training schemes that certify achievement of a coherent set of skills and knowledge and that have evidence of need by industry, employers, [community of people] and/

or the community.”

UNESCO (Oliver & UNESCO, 2022) has identified the following features in MCs, which give rise to multiple affordances, namely:

- being human-centric,
- promoting both equity (United National Sustainable Development Goal Four) and digital transformation/aiming to bridge the digital divide,
- possessing diversity in stakeholders,
- forming an agreement on the scope and definition of MCs,
- agreeing on how to quality-assure, recognise, regulate, and incentivise them,
- being flexible, portable, transferable, and transparent,
- having agreed on learning outcomes/achieved competencies and
- ensuring they are not over-regulated.

I would suggest that an MC be preferred over a *digital badge*, the latter being a broader term. By their very nature and definition, MCs are *small* and *assessed* parcels of learning. In my learning journey, I realised that MCs can be distinguished from badges through the lens of assessment (Hanshaw, 2023); writers invariably refer to digital badges when there is little or no assessment (Grant, 2016).

There is now general agreement, if at times consternation, that MCs can be *stacked*. Lockley et al. (2016) argue that pegging MCs to existing frameworks can be cumbersome. Gibson et al. (2016) put forward that stackable MCs are a new means of identifying skills, experience and knowledge and that there is the possibility to use badges in all three stages of the learning journey: paths *into* learning, paths *during* learning, and *lifelong* learning pathways which the European Commission (2020) comments on the lack of consistency and standardisation in MCg. Thus, MCs can be bundled together to create a series of awards, potentially culminating in and triggering a more significant award. For example, a series of 15-credit MCs could be stacked to form 60 credits, thereby triggering an exit award of a 60-credit certificate or 120-credit diploma at the required and achieved level. A capstone assessment is likely an excellent strategy to double-check that learning outcomes/achieved competencies have been demonstrated before triggering the stacked exit award. Theoretically, high-level MCs could be stacked to form the credits necessary to trigger a degree-level qualification at the undergraduate or postgraduate level. However, as the literature will reveal, this affordance has its critics.

## Affordances

Affordance can be defined as what is furnished or provided by an agent to a party or parties. For example, numerous environmental affordances exist for animals (Gibson, 1979). These can be for good or bad and are complementary: the environment complements the animal and vice versa (Gibson, 1979). In this context, the agent is the MC and the party, or parties, are the learners.

McGreal and Olcott (2022) outline the multiple affordances of MCs within higher education, using the case study of Deakin University (DeakinCo, 2017), which considers how an organisation might achieve a competitive advantage in their strategic deployment of MCs. They posit that short courses that lead to micro-credentials can afford employees and employers flexibility and “just-in-time training... empowering employees to upskill, learning how to function in emerging new critical areas for an industry” (McGreal & Olcott, 2022, p. 5). Emergent knowledge or skills, or those required urgently, need an urgent and manageable response not provided by a traditional degree. MCs can be deployed in this space quickly and achievably, resulting in timely and empowering success.

MCs have multiple affordances in an HE/VET context. Gibson et al. (2016) consider MCs for supporting learning journeys: bringing visibility and transparency to the learning, teaching and assessment journey, illuminating the affordances of learning to stakeholders, and providing a new means of identifying skills, experience and knowledge “through an open, transferable, stackable technology framework” (p. 115). They also consider the importance of MCs in sustaining life-long learning and argue that MCs enable the institution to leverage the building of professional networks.

Wilson et al. (2016) consider how MCs are often contrasted with degrees: MCs, unlike degrees, may be issued by employers and professional bodies and accessed flexibly by learners, and they help institutions move away from a seat-time model towards a competency-based curriculum, by which they mean MCs are “disruptive innovation” (p. 164). Lockley et al. (2016) consider them a disruptive technology (questioning the status quo). A European Commission report on MCs, *Micro-credentials in the EU and Global* (2020), finds that there is disagreement amongst experts as to whether educational institutions will get disrupted by companies offering MCs. Wilson et al. (2016) reflect on how MCs in the university system are like David and Goliath: David is the upstart MC, Goliath the institution. The story talks of Goliath being fierce and bigger than others, with a sword and spear to attack and a large shield to defend. This is how one might imagine the neoliberal university: its size, ease of defence with lawyers for a shield, and sharp tools for attack. David, conversely, had only his faith. Coleman and Johnson (2016) endorse this David and Goliath analogy, arguing that MCs have provided HE with the ability to recognise detailed aspects of learning; they enable the endorsement of competencies, capabilities and skills: those that go unrecognised within the traditional academy or in the transcript.

## Disruption

Disruptive innovations create footholds in markets where no market existed, turning the non-consumer into the consumer. However, their success for mainstream consumers is quality-dependent: they do not become popular until they possess sufficient quality to satisfy the mainstream consumer (Christensen et al., 2015). This suggests MCs, if disruptors, might provide open access opportunities for learners to get a foothold in post-secondary education where before, no such opportunity existed, turning the non-learner into a learner. The quality of the MC offering will vary between awarding organisations. However, one would have thought that if a reputable institution were administering the MC, perceptions of quality should be satisfied, and expectations should be met. The debate is whether MCs as disruptors are unseating or augmenting the traditional university credential. In the later part of this article, I shall consider whether MCs are inferior to traditional credentials or have the potential to unseat or contribute to unseating. For now, it is worth noting what Kumaraswamy et al. (2018) highlight – “Many years ago, Peter Drucker noted, ‘The greatest danger in times of turbulence is not turbulence itself, but to act with yesterday’s logic’” (p. 19).

A key question is whether *disruption* is positive in creating opportunity where none existed or is harmful in displacing or replacing something qualitatively inferior. Replacement is a subjective term: who or what is being replaced? If the traditional degree is being replaced by MCs, that may displace the traditional academic. However, it may afford opportunities to learners. Is raging against the disruptive element of MCs an example of acting with yesterday’s logic against an innovation? One person’s negative disruption may be a positive innovation for another.

### Micro-credentials as negative disruptors

MCs have their critics, including those who regard them as reductionist and a threat to traditional education. Ralston (2020) is one of MC’s greatest critics, calling MCg “dangerously reductivist” (p. 95) and “a moral hazard” (p. 96):

It reduces higher learning to a list of hard skills and technical competencies that bolster employer workforce development and heighten employees’ earning potential. Soft skills and human competencies to, for instance, ‘learn to learn’ are arbitrarily excluded from micro-credential curricula (Ralston, 2020, p. 95).

MCg can enhance career development and personal growth (Grugulis & Vincent, 2009), something that Ralston concedes. However, Ralston suggests that “Micro-credentialing contributes to the decline of the traditional degree. It paves the way for the total substitution of degree programs with micro-credentials” (p. 95). Has any traditional degree programme ever been substituted by MCs in global higher or vocational education? Not to my knowledge. That is not to say that MCs might not go on to replace some traditional university awards. However, this has not happened to date.

Artificial Intelligence (AI) has been posited as replacing teachers in learning, teaching, and assessment. However, the reality is that AI likely augments and adds value to the traditional practitioner (Crawford et al., 2023). AI cannot replicate the human touch in learning, and it is essential to recognise the crucial role humans play in education and navigating changes in technology (Osamor, 2024). Generative AI can bring advantages in terms of efficiency for both educators and learners (Rudolph et al., 2023). Indeed, AI “presents the opportunity to re-emphasize that a university can serve the common good and shift towards a better future” (Popenici et al., 2023a, p. 103), fostering respect for learners and academics as we move towards a common goal. However, we are likely quite unprepared for AI, which will force institutions to ask themselves what they are doing: humans might be removed from the learning and assessment processes altogether, resulting in no one learning anything, and technology without human morality poses some threat (Popenici et al., 2023b). Whether it poses an existential threat remains to be seen, although perhaps not seen for very long, by us anyway, should that be the case.

MCs could be associated with the perhaps lesser threat of the evils of neo-liberalism and market-driven education as “microcredentialing generates a consistent stream of revenue through planned obsolescence, perpetual servicing, and moral hazard” (Ralston, 2020, p. 17). However, Desmarchelier and Cary (2022) provide an interesting critique of Ralston’s position. Do MCs pander to the neoliberal ideology ornate? They ask:

What is sacred about a traditional degree structure? We see the undergraduate/postgraduate degree structure as firmly embedded globally in neoliberal education systems that require the expenditure of (usually large) amounts of money from varying mixtures of private individual and public government sources (Desmarchelier & Cary, 2022, p. 5).

Desmarchelier and Cary (2022) argue that the demands of neoliberalism are strongly present in traditional degree programmes: “Universities are increasingly beholden to economic imperatives, and efficient delivery of learning is a goal at most institutions” (p. 5). Andrew (2023a) asserts that “universities are increasingly managed and neoliberalised, corporatising and commercialising” (p. 18) and investing in real estate empire-building rather than funding academic positions (Andrew, 2023b). This suggests that university leadership is already strongly influenced by the demands of neoliberalism, with MCs having little influence on the property empire ambitions of university boards.

These statements resonate with anyone who has worked in the academy. Why is MCg the target for these accusations, rather than “the fastidious micromanager, marked by an inward deep feel of failure” (Andrew, 2023a, p. 18)? Or the endless marketing missions overseas to recruit international dollars? The competition for research funding or tenure may be more brutal than the effects of MCg upon the academy and the academic.

Does academic snobbery play a part in the disdain towards MCg from some corridors of the academy? At the heart of this, perhaps, is that MCg is usually considered vocational, which is “strongly perceived” as being “second choice” (Keevy, 2020, p. 1). This view is only heightened by those who consider the academy sacred. Ralston (2020, p.12) notes:

What is lost in the conversion of higher education to a microcredentialing delivery system is the rich educational experience whereby teacher-scholars share new vocabularies, culture and dispositions... in an ongoing and mutually edifying conversation. Also abandoned is the higher purpose of education, namely, to serve society at large, not simply corporations and industry.

Is HE being *converted* into MCg? Why can learners and teachers not share essential learning moments in a micro-learning environment? Are these ‘edifying conversations’ not instances of micro-learning themselves? I would argue that Ralston’s position results from a philosophical aversion to MCs as being negatively disruptive, threatening to displace or replace him as a traditional academic rather than affording opportunity for micro-learning when, perhaps, learner and teacher are in the same space but not necessarily the same place, for example, online.

Do MCs generate revenue at the cost of ethical responsibility? MCs contribute to the decline of the degree (Kazin & Clerkin, 2018) and undermine the very mission of HE by promoting efficiency and profitability (Ralston, 2020). However, do they pose as significant a threat as using university money to develop property empires rather than fund academic positions? Ralston (2020) posits: “Administrators who invested in microcredentialing as a revenue generator will sometimes have to shirk their ethical duty to act in the best interests of students in order to maximize profits” (p. 17). Why is this truer for MCg than any other HE or VET learning vehicle? Managers instruct administrators, and it is the managers who drive the fulfilment of financial and recruitment targets as well as pass rates for programmes at all levels of HE and VET: “Institutions are more interested in getting students ‘over the line’ in a timely fashion than in facilitating opportunities for authentic excellence” (Andrew, 2024, p. 1). It may perhaps resonate with some readers that it is commonplace for recruiting universities in the so-called developed world to occasionally abandon much of their ethics in respect of entry requirements to satisfy recruitment targets for degree and vocational programmes in the highly competitive and lucrative international student recruitment market, at all levels of the HE and VET sectors. It could equally be argued that neoliberal demands and the ‘shirking of ethical duty’ were present in education long before MCs arrived.

Ralston asserts that MCg’s “focus on vocational education allies it with vested industrial and corporate interests... For Marxists, credentialism suggests bourgeois values that, when pursued by proletariat members, generate a version of false consciousness” (Ralston, 2020, p. 18). Do MCs align themselves with such vested interests more than traditional programmes? To graduate from Harvard, Yale,

Princeton, MIT, Oxford, or Cambridge is undoubtedly the epitome of academic power and often privilege, vested self-interest, affording opportunity for the few at perhaps the expense of the many: not so much bourgeois, middle-class, as elite, upper-class. There is undoubtedly greater 'false consciousness' in the elevation of the so-called elite into positions of power and privilege than in widening participation in the education of the 'proletariat', which Ralston appears to argue against.

The contention here seems to be that MCs are less than whole by their micro nature. However, degree programmes are already unbundled into modules or courses, lectures and seminars, assessments and tutorials. None of these moments in the learning journey educates the whole person all at once. Therefore, the claim that MCs are less than whole is undoubtedly a lame claim, as neither is the traditional degree in its parts.

Desmarchelier and Cary (2022) grapple with this: "A consideration when unbundling learning occurs is the maintenance of rigour and quality and the cohesiveness of the learning offered. To produce quality-assured micro-credentials from existing material requires significant development" (p. 7). This has a significant development cost, and while Ralston is right that unbundling traditional degrees is a source of material for MCs, such MCs can have an efficacious effect on learners' learning if undertaken with judicious rigour (Desmarchelier & Cary, 2022).

Ralston (2020) argues that "microcredentialing does not liberate learners' potentialities or meet the needs of lifelong learners" (p. 96). Desmarchelier and Cary (2022) "reject the generalised assumption that micro-credentials only pertain to industry-specific skills and competencies and have demonstrated how universities respond in ways that develop learning-to-learn proclivities." (p. 7). There is undoubtedly a place for MCg within HE and VET to develop discrete skills and promote life-long learning through discrete parcels of learning. However, Lewis and Lodge (2016) argue that the:

A reductive MC approach to professional practice generation and CPD should be used for lower-order or vocational skills that can be simply noticed. The level of granularity does not provide the nuances required for higher-order processing and the subtleties of knowing, being, doing, and valuing. A more holistic approach is required for the complexities of uncertain workplaces (Online).

Grant (2016) states the opposite: MCs are particularly relevant to enhancing nuanced understanding and allowing more transparency and an evidence base. A mature debate indeed leads us to a familiar place, in the realisation that whether any credential has higher value is determined by how it is administered, delivered, assessed, moderated, awarded, and quality assured, as well as to the, sadly, familiar space of equating the vocational with the lower order (Lewis & Lodge, 2016). The inherent snobbery towards, or rather against, the vocational education and training sector (Meade & Feldman, 1966) will continue to have a deleterious effect on HE as it excludes the vocational from the HE table, creates a barrier for learners and learning, creating an unnecessary

divide that need not exist, hindering opportunity to HE to broaden and expand its offering, hampering VET by impeding integration within HE, and discouraging learners from pursuing a vocational route.

Wheelahan and Moodie (2022) are also vociferous critics of MCg:

Rather than presenting new opportunities for social inclusion and access to education, they contribute to the privatisation of education by unbundling the curriculum and blurring the line between public and private provision in higher education (p. 1288).

Public-private partnerships are not new (Breton & Lambert, 2003). If MCg contributes to new partnerships, why is that a problem, and why is that preventing new opportunities for social inclusion and access? Wheelahan and Moodie (2022) assert: "They [MCs] accelerate the transfer of the costs of employment preparation, induction, and progression from governments and employers to individuals" (p. 1279). That is an interesting interpretation of individuals choosing to up-skill or re-skill for employability or enhancement. Do degree programmes not transfer costs from governments and employers to individuals? Governments, decreasingly, and employers rarely fund degree programmes contemporaneously; individuals fund themselves. However, should degree programmes be dismissed because they transfer costs from the government and employer to the learner? Perhaps they should be critiqued on this basis, if not dismissed as such.

Wheelahan and Moodie (2022) themselves critique:

Microcredentials contribute to 'disciplining' higher education in two ways: first by building tighter links between higher education and workplace requirements (rather than whole occupations), and through ensuring universities are more 'responsive' to employer demands in a competitive market crowded with other types of providers (p. 1279).

I consider both of these affordances as positive drivers of MCg: to have tighter links between HE and industry and to be more responsive to the needs of employers, demonstrate accountability to the needs of society as a whole, employers in particular, and learners specifically. By 'whole occupations,' do the authors mean doctors, dentists, lawyers, and politicians? MCg can contribute to universities and vocational institutions being more useful, responsive, accessible, affordable, and less overbearing as a learning proposition. However, the authors dismiss this, arguing that MCs "are gig credentials for the gig economy" (Wheelahan & Moodie, 2022, p. 1281) and do not challenge the status quo, and those not attending elite institutions must still second-guess the labour market when it comes to upskilling.

It is hard to imagine that snobbery can ever be eradicated in education, and there are those whose interests it serves who would never wish to do so. However, as for the proletariat, second-guessing and upskilling may be all we can hope

for, so why deny us that? MCs provide a vehicle to improve ourselves, whoever we are, stand out, and go further, wherever we are, "at the right time and for the right job" (Wheelahan & Moodie, 2022, p. 1281).

Those that oppose MCs are often type-cast as moth-balled die-hards (Wheelahan & Moodie (2022):

Opponents of microcredentials are cast as those who wish to maintain higher education as an ivory tower and support elite structures of higher education, who are conservatives resistant to change and who deny any role for higher education in supporting people to gain credentials they need for a meaningful career (p. 1281).

MCs are perhaps not the antidote to elitism in higher education; its opponents are not necessarily conservatives in ivory towers, and degree programmes can and do contribute to meaningful careers. However, there is no doubt that an MC is more accessible than a whole degree (even when a degree is divided into years of study) for cost and time reasons (Tehan & Cash, 2020). They may not revolutionise education for the better; indeed, they have not. However, they may make it better for many by creating vocational and higher education opportunities. After all, MCs cannot only be for the lucky few (Oliver & UNESCO, 2022).

Wheelahan and Moodie (2022) argue that rather than investing in MCs, "progressive, democratic societies should seek to ensure that all members of society have access to a meaningful qualification that has value in the labour market and society more broadly" and to enable individuals to live valuable lives (p. 1279). Rather than focusing on this to the exclusion of MCs, I argue that MCs can be a conduit to achieving this, given their ability to afford access and equity on the grounds of affordability and time/financial/academic/cognitive manageability. MCs can be a powerful tool for enabling equity, access, and participation. Likely, society does not deny people experiencing homelessness a meal because they have not yet provided them with access to a job and a safe place to sleep. Higher and vocational education should not deny learners and would-be learners access to MCs just because not all societies have access to education; contrary, it may just be the antidote, if not the panacea. That small gift to a homeless person may go some way to alleviating their poverty, or at least sustaining them today; MCs can go some way to alleviate intellectual and skill-set poverty and increase and expedite access to new vocations and new intellectual spaces. That, indeed, is something Ralston et al. (2020) can get behind.

Marshall (2010) reminds us:

Disruptive change is problematic for dominant organisations as the natural tendency is to protect existing structures and activities, particularly when those are currently seen as successful (p. 181).

This supports the view that innovation displaces traditional structures. However, with MCs, can they not augment and complement? I shall consider this below.

## Micro-credentials as agency

The term 'agency' can mean "action or intervention producing a particular effect," for example, "canals carved by the agency of running water" (Bab. Ia, 2023). Here, we do not consider carving canals but knowledge, skills, and more fulfilling lives by the agency of MCs.

It is not only learners but industry or society as a whole that can benefit from nimble, rapidly deployed MCs. By their very nature, MCs enable HE and VET to respond quickly to individuals' educational needs, enable learners to upskill and find "more meaningful and lucrative participation in the workforce", as well as afford "dipping their toes in the water" for further traditional higher education study with many universities promoting MCs as stackable for credit to provide pathways into macro qualifications" (Desmarchelier & Cary, 2022, p. 6). Emergent skills and knowledge need to be acted on now, not when the next group of first-year students graduate in three or four years. Rather than threatening the traditional degree, this is a different beast altogether: put crudely, this is the corner shop, not the department store. When one needs something urgently, one goes to the corner shop. Further, the idea of 'dipping your toes in the water' is powerful. A learner can sample computing with an MC to see if it is what they want, without needing to get part-way through a lengthy degree programme only to discover they would instead stack shelves than work in IT.

Turning to life-long and life-wide learning, contrary to the idea that only an arts education can afford the beauty of intellectual exploration, learners may use MCs to further develop themselves in any area of interest, need, or desire, such as numeracy, literacy, family health and well-being, writing, or participation in lobbying and the democratic process, whereby contributing to life-wide learning (Desmarchelier & Cary, 2022). MCs can, therefore, afford lifelong learning and life-wide learning that are available where required or desired. The idea of MCs being deployed better to enable participation in activism or the democratic process is robust. To those who take aim at the verb 'invest in', this is precisely what most learners do with a traditional degree: it is not free. It is an investment in cultural capital. However, arguably, it is increasingly a form of taxation as student debt can remain with learners for life. At the same time, the degree might become a financial ball and chain; an MC could, dare we dream, be a means of emancipation, though, as we have seen, this view has critics, as it should. However, what if we should dare to dream?

Arguably, the most potent agency of MCg is in its affording equity, access, and participation in higher education. As Desmarchelier and Cary (2022) note:

Rather than forcing potential students into lengthy, expensive degrees, micro-credentialed offerings that can be accessed as either lifelong or life-wide learning needs that arise, mean more and cheaper access to education than previously available (p. 8).

MCs can, therefore, promote equity, access, and participation in HE and VET through affordability and the more realistic undertaking of smaller chunks of learning, contributing to

lifelong and life-wide learning. The European Commission (2020) observes that “affordability has become one of the drivers for the growing use of micro-credentials. Particularly, in the US, there is evidence that people are starting to question the benefits of paying for traditional qualifications that may not prepare them for the new digital society” (p. 40). Varadarajan et al. (2023) take up this point: “Financial barriers to undertaking microcredentials do not comprise significant barriers for learners. Higher education institutions can offer microcredentials in smaller units at lower costs than full-term tuition fees” (p. 14).

The term *tuition fees* reminds us that higher and vocational education are already vultures at the neoliberal sacrificial altar, dividing up the spoils. Rather than accelerating the consumption of dollars and barring access to those who cannot afford to sit at the table, MCs can ameliorate this by enabling more people to learn more about it. Time and cost are two of the greatest barriers to HE (Tehan & Cash, 2020); MCs reduce them both. Fear of failure and a perceived lack of success are two of the greatest barriers in HE and VET (Hanshaw, 2023); MCs put learning into more manageable chunks, reducing that fear and enabling learners to see success in incremental steps.

MCs have the potential to alleviate student drop-out rates (Pirkkalainen et al., 2023). They also put learning in the hands of the learner, who can more readily decide what they learn and when (Hanshaw, 2023), which will logically contribute positively to retention rates on programmes of study. These are powerful, positive affordances.

MCs can enable the dissemination of learning and credentialing on a global scale. Desmarchelier and Cary (2022) argue that MCs enable learning to be internationalised in an unprecedented manner:

The digital allows for global access to education in a way never seen before. A course can be offered by an Australian university and have participants from South America, Asia, and Europe, making for an enriched learning environment for students (p. 9).

Though this is equally true of any, or most other, packages of learning and assessment, the affordance of MCs leveraging equity, access, and participation within HE and VET makes global access to such learning opportunities even more powerful: again, more people can learn more about more. It can also enable open access, for example, the Open Education Resource Foundation (OERu), based at Otago Polytechnic in Aotearoa, New Zealand, which is making a landmark contribution in providing open access resources to learners and scholars:

One of the most innovative organizations in the world to combine online learning, OER and open systems across digital formats with a diverse system of micro-credentialing. OERU offers a range of short courses and seminars for non-credit that are stackable together into traditional credentials with partner universities (McGreal & Olcott, 2022, p. 12).

This stackability further empowers learners, as they can use MCs to contribute to or even trigger awards within the traditional HE or VET systems in a more affordable, time-friendly, and less daunting manner than the traditional chunk of learning that is a degree. This brings us to another powerful affordance of MCg: integration into the university curricula.

The integration of MCS within the curriculum is a complex and contentious issue. McGreal and Olcott (2022) argue that integration can “make them [MCs] easy to use with clear validation metrics, and, in this way, make micro-credentials a value-added benefit for all stakeholders” (p. 6). This is a clever way to deploy pre-established and robust quality assurance mechanisms to achieve economy of scope to the benefit of the institution, employers/industry, and learners by creating targeted discrete parcels of learning that can be incorporated into more extensive programmes, or through stacking, can trigger larger awards, with capstone assessments for added rigour in high stakes credentials: where the institution or industry require reassurance that learners are sufficiently capable in the target knowledge or skill-set, whether traditional (e.g. English for Coastguards) or emergent (e.g. how to administer the COVID vaccine).

I have already considered how MCs can enable emergent, urgent knowledge and skills to be developed and recognised in a more expeditious manner than the traditional degree. However, this new learning can then be integrated into the curriculum. This can provide added value in the exchange of knowledge. Further, by redesigning curricula into a series of MCs, where an institution is seeing a low rate of admission or high dropout rates, MCg the curriculum could enable enhanced learner success (Hanshaw, 2023). Thus, “embedding micro-credentials within the curriculum has the potential to affect how students understand their social and cultural capital” (Pollard & Vincent, 2022, p. 852). This could also be done by integrating the students' knowledge and skills into the curriculum, not just their lecturers. Students could be encouraged to redeploy their ideas into MCs to the benefit of many. This is harnessing expertise as we have never seen before. Not to forget the expertise within faculty, however:

Micro-credentialing represents a potential seismic shift in the global landscape of higher education. Most institutions will have pockets of highly innovative learning and teaching practice driven by committed academic staff. To make micro-credentialing successful, these need to be harnessed and directed at a whole of institution level (Demarchelier & Cary, 2002, p. 9).

However, to harness this potential, HE and VET sectors need to stop arguing over whether MCs are an existential threat when they have proved themselves not to be and engage in the serious business of making them work:

Micro-credentials [need to] become more widely accepted and standardised, meaning that national strategies would need to be strengthened, micro-credentials aligned with national qualifications systems and the policies for the common assessment strategies of micro-credentials outlined (Pirkkalainen et al., 2023, p. 43).

This is not such a tall order. Traditional degree programmes are integrated or stand-alone, and reputable institutions seem to succeed in turning out quality graduates. Why cannot MCs do the same? Boud and Jorre De St Jorre (2021) warn us not to deploy MCs only in unbundling existing qualifications “until these are reformed to be transparent in meeting the minimum standards of achievement required for each designated outcome. Without this, flaws associated with macro-credentials will inadvertently undermine micro-credentials” (p. 19).

Unbundling existing qualifications need not be an issue – since COVID-19, the unbundling of HE has already commenced (Varadarajan et al., 2023); if minimum standards are agreed upon and transparent if achieved competencies/ learning outcomes are accurately and clearly assessed and moderated, in short, if there is capability in academic practice. Perhaps the problem can become its solution: rather than MCs being reductionist, they can inspire and build capability in manageable chunks of learning, teaching, and assessment practice across institutions within HE and VET. This task is more manageable for practitioners and learners than grappling with a whole programme all at once. Is grappling with the complexities of MCs within HE and VET worthwhile? Ask a pertinent question: Should institutions engage in MCs at all?

The critical strategic reset question for university leaders is not how we engage in micro-credentials. The first question is, should we engage at all, or to what degree? Does this trend align with the institutional mission, the programmatic strengths of academic programmes, and the increasing importance of good judgment in strategically allocating institutional resources? Indeed, like online learning and open education, one can effectively make the case for all institutions to be involved in workforce and economic development; however, at the end of the day, the related question is how involved? (p. 15).

The answer to this question depends on *whose* interests one’s institutional strategy serves. Is it the interests of the registrars, the programme leaders, IT, and Finance that bulk at the amount of work and money involved in integrating stackable MCs into the current curriculum, or in using the current curriculum to create a stack of them, then assessing them, moderating them, reviewing them, recording them, and creating a repository that learners can access and even employers to display or validate them? Alternatively, is it in the interests of the learners, the would-be learners, the never-would-have-been learners, who could seize the opportunity to undertake manageable, less frightening, affordable, time-friendly, stackable, transferable, validated, assessed chunks of learning and assessment of their choosing, that complement the traditional offering with life-long and life-wide learning, and do not replace it? Widening Participation (WP) has long been a central agenda of institutions I have worked for in the United Kingdom and New Zealand – in theory. MCs afford the potential to revolutionise the WP agenda in practice. Therefore, the overarching question is, are we up to the challenge to make a real difference in the lives of many by making our degree offerings more

nimble to change and more available, affording change for the betterment of us all, not just the lucky few (Oliver & UNESCO, 2022).

## Conclusion

In this article, I have briefly defined MCs as small chunks of assessed learning that are developed, administered, and awarded with the quality assurance that one would expect from an institution of high standing. I have reflected on some of MCs’ most vociferous critics: MCg dumbs down learning; is an unethical revenue generator; does (or does not) enable higher order processing; contributes to the privatisation of the academy; constrains institutions, and rather than liberating learners, forces them into paying for what employers should be paying for – upskilling or re-skilling.

I have also uncovered some powerful affordances of MCs as positive agents: enabling quick responses to changing needs of individuals, organisations, or societies; for learners to sample a subject area or field of practice with an MC without committing to a lengthy and costly degree; provide life-wide as well as life-long opportunities for personal and societal growth; to put the decision of what to study and when more in the hands of the learner rather than the institution; afford equity, access, and participation by virtue of MCs reducing time and money constraints; promote new learning that can be integrated into the curricula; alleviate a fear of failure or lack of success by redesigning existing qualifications into manageable chunks whereby improving recruitment onto programmes and retention rates within programmes; enable greater access to education for more people in more places, who can come together and learn together; provide a student voice to knowledge generation and dissemination by integrating the learning of learners into the curricula; and act as a conduit to capability building within the academy in the development of curricula and quality assurance in manageable chunks of subject matter and practice.

Finally, I am struck by two things: first, the positive affordances of MCg appear to significantly outweigh the critics’ claims of negative affordances in size and number. This is following my honest attempt to research the literature without bias in this space.

Secondly, the ideas of those who argue that MCs are a disruptive force for evil are just that: ideas. There is little evidence to suggest the claims to be true. We can claim that MCs are a “moral hazard” (Ralston, 2020, p. 96). However, there is no evidence of one learner or academic being displaced or harmed by them. It can be argued that MCs develop higher-order processing skills or do not; however, no evidence supports such claims. However, the examples of MCs being an innovative force for good are logical. We do not need evidence to understand that an MC is cheaper and more accessible than a degree (however, evidence can be provided); it can respond faster to a changing environment than a degree; an MC can provide meaningful learning in areas of personal growth and development; it puts learning decisions, what to learn and when, into the hands of the

learner, not just the institution; it is an excellent way to 'suck it and see' before enrolling on a degree programme or macro qualification; the learnings from MC knowledge generation and skill acquisition can be integrated into the curricula; more people can come together to learn what they want and when they want; and MCs make academic programmes and academic developments more manageable in size and task. This is logical and common sense. As positive disruptors, MCs provide a foothold into learning, making learners of non-learners or would-be learners. They have yet to displace the academy or the academic, that is, to be negatively disruptive, except for the debate on their potential disruption, which distracts from the more critical task of educational enhancement, especially post-COVID, where many institutions feel disjointed.

Therefore, moving forward, I recommend that we stop disappearing down the rabbit hole of debating whether MCs are good or bad, whether they are best seen as vocational or HE, but look at how we might harness them for the betterment of the many:

There is strong hope that micro-credentials can advance the equity agenda, bringing accessible and affordable focused learning and skill building to vulnerable communities, enabling achievement of the United Nations Sustainable Development Goal 4 (Quality education) (Desmarchelier & Cary, 2022, p. 8).

Let us travel in hope.

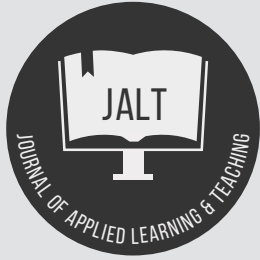
## References

- Andrew, M. B. (2023a). Come to the cabaret: Voices from the modern university. *Journal of Applied Learning and Teaching*, 6(2), 17-27. <https://doi.org/10.37074/jalt.2023.6.2.19>
- Andrew, M. B. (2023b). Neo-neoliberalist capitalism, intensification by stealth and campus real estate in the modern university in Aotearoa/New Zealand. *Journal of Applied Learning and Teaching*, 6(2), 393-401. <https://doi.org/10.37074/jalt.2023.6.2.16>
- Andrew, M. B. (2024). 'Just get them over the line': Neoliberalism and the execution of 'excellence'. *Journal of Applied Learning and Teaching*, 7(1), 1-11. <https://journals.sfu.ca/jalt/index.php/jalt/article/view/1637>
- Bab.la. (2023). *Agency*. In en.bab.la.com. <https://en.bab.la/dictionary/english/agency>.
- Boud, D., & Jorre De St Jorre, T. (2021). The move to micro-credentials exposes the deficiencies of existing credentials. *Journal of Teaching and Learning for Graduate Employability*, 12(1), 18-20. <https://doi.org/10.21153/jtlge2021vol12no1art1023>
- Breton, G., & Lambert, M. (2003). *Universities and globalization: Private linkages, public trust*. UNESCO Publishing.
- Christensen, C. M., Raynor, M., & McDonald, R. (2015, December). What is disruptive innovation? *Harvard Business Review*. <https://hbr.org/2015/12/what-is-disruptive-innovation>
- Coleman, K. S., & Johnson, K. V. (2016). Badge claims: Creativity, evidence and the curated learning journey. In D. Ifenthaler, N. Bellin-Mularski, & D. K. Mah (Eds.), *Foundation of digital badges and micro-credentials* (pp. 369-387). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15425-1\\_20](https://doi.org/10.1007/978-3-319-15425-1_20)
- Crawford, J., Vallis, C., Yang, J., Fitzgerald, R., O'Dea, C., & Cowling, M. (2023). Editorial: Artificial intelligence is awesome, but good teaching should always come first. *Journal of University Teaching and Learning Practice*, 20(7). <https://doi.org/10.53761/1.20.7.01>
- Deakin Co. (2017). *What are micro-credentials and how can they benefit both businesses and employees?* <https://www.deakinco.com/media-centre/article/Benefits-of-micro-credentials-for-business-and-employees>
- Desmarchelier, R., & Cary, L. J. (2022). Toward just and equitable micro-credentials: An Australian perspective. *International Journal of Educational Technology in Higher Education*, 19(1), 25. <https://doi.org/10.1186/s41239-022-00332-y>
- European Commission. (2020). *European Commission expert report on micro-credentials*. European Commission. [https://microcredentials.eu/wp-content/uploads/sites/20/2021/01/European-Commission-expert-report-on-microcredentials\\_Microbol\\_27.01.pdf](https://microcredentials.eu/wp-content/uploads/sites/20/2021/01/European-Commission-expert-report-on-microcredentials_Microbol_27.01.pdf)
- Gibson, D., Coleman, K., & Irving, L. (2016). Learning journeys in higher education: Designing digital pathways badges for learning, motivation and assessment. In D. Ifenthaler, N. Bellin-Mularski, & D. K. Mah (Eds.), *Foundation of digital badges and micro-credentials* (pp. 115-138). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15425-1\\_7](https://doi.org/10.1007/978-3-319-15425-1_7)
- Gibson, J. J. (1979). *The theory of affordances*. Boston: Houghton Mifflin. [https://monoskop.org/images/c/c6/Gibson\\_James\\_J\\_1977\\_1979\\_The\\_Theory\\_of\\_Affordances.pdf](https://monoskop.org/images/c/c6/Gibson_James_J_1977_1979_The_Theory_of_Affordances.pdf)
- Grant, S. (2016). Building collective belief in badges: Designing trust networks. In D. Ifenthaler, N. Bellin-Mularski, & D. K. Mah (Eds.), *Foundation of digital badges and micro-credentials* (pp. 97-114). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15425-1\\_6](https://doi.org/10.1007/978-3-319-15425-1_6)
- Grugulis, I., & Vincent, S. (2009). Whose skill is it anyway?: 'Soft' skills and polarization. *Work, Employment and Society*, 23(40). <https://doi.org/10.1177/0950017009344862>
- Hanshaw, J. (2023). Leveraging, and not leveraging, micro-credentials – and a global pandemic. *Scope: Contemporary Research Topics (Work-based Learning)*, 5, 46-56. <https://doi.org/10.34074/scop.6005004>
- Kazin, C., & Clerkin, K. M. (2018). *The potential and*



- limitations of microcredentials. Service Members Opportunities Colleges. [https://supportsystem.livehelpnow.net/resources/23351/Potential%20and%20Limitations%20of%20Microcredentials%20FINAL\\_SEPT%202018.pdf](https://supportsystem.livehelpnow.net/resources/23351/Potential%20and%20Limitations%20of%20Microcredentials%20FINAL_SEPT%202018.pdf)
- Keevy, J. (2020). *(Un)recognising learning in our digital age*. Johannesburg: JET Education Services. [https://www.academia.edu/41515579/\\_Un\\_recognising\\_learning\\_in\\_our\\_digital\\_age\\_Working\\_Draft](https://www.academia.edu/41515579/_Un_recognising_learning_in_our_digital_age_Working_Draft)
- Kumaraswamy, A., Garud, R. & Ansari, S. (2018). Perspectives on disruptive innovations. *Journal of Management Studies*, 55(7), 1025-1042. <https://doi.org/10.1111/joms.12399>
- Lewis, M. J., & Lodge, J. M. (2016). Keep calm and credential on: Linking learning, life and work practices in a complex world. In D. Ifenthaler, N. Bellin-Mularski, & D. K. Mah (Eds.), *Foundation of digital badges and micro-credentials* (pp. 41–54). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15425-1\\_3](https://doi.org/10.1007/978-3-319-15425-1_3)
- Lockley, A., Derryberry, A., & West, D. (2016). Drivers, affordances and challenges of digital badges. In D. Ifenthaler, N. Bellin-Mularski, & D. K. Mah (Eds.), *Foundation of digital badges and micro-credentials* (pp. 55–70). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15425-1\\_4](https://doi.org/10.1007/978-3-319-15425-1_4)
- Marshall, S. (2010). Change, technology and higher education: Are universities capable of organisational change? *Research in Learning Technology*, 18(3), 179–192. <https://doi.org/10.1080/09687769.2010.529107>
- McGreal, R., & Olcott, D. (2022). A strategic reset: Micro-credentials for higher education leaders. *Smart Learning Environments*, 9(1), 9. <https://doi.org/10.1186/s40561-022-00190-1>
- Meade, E. J., & Feldman, M. J. (1966). Vocational education: Its place and its process. *The Journal of Human Resources*, 1(1), 70–74. <https://doi.org/10.2307/145017>
- New Zealand Qualifications Authority. (2020). *Microcredentials*. <https://www.nzqa.govt.nz/providers-partners/approval-accreditation-and-registration/micro-credentials/>
- Oliver, B., & UNESCO. (2022). *Towards a common definition of micro-credentials*. <https://unesdoc.unesco.org/ark:/48223/pf0000381668>
- Osamor, A. (2024). Rise of the robots: What it means for educators. *Journal of Applied Learning and Teaching*, 7(1), 1-11. <https://doi.org/10.37074/jalt.2024.7.1.16>
- Pirkkalainen, H., Sood, I., Padron Napoles, C., Kukkonen, A., & Camilleri, A. (2023). How might micro-credentials influence institutions and empower learners in higher education? *Educational Research*, 65(1), 40–63. <https://doi.org/10.1080/00131881.2022.2157302>
- Pollard, V., & Vincent, A. (2022). Micro-credentials: A postdigital counternarrative. *Postdigital Science and Education*, 4(3), 843–859. <https://doi.org/10.1007/s42438-022-00311-6>
- Popenici, S., Catalano, H., Mestic, G., & Ani-Rus, A. (2023a). A systematic review of the artificial intelligence implications in shaping the future of higher education. *Educatia*, 21(26). <http://dx.doi.org/10.24193/ed21.2023.26.11>
- Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023b). A critical perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning and Teaching*, 6(2), 311-331. <https://doi.org/10.37074/jalt.2023.6.2.5>
- Ralston, S. J. (2020). Higher education's microcredentialing craze: A postdigital-Deweyan critique. *Postdigital Science and Education*, 3, 83-101. <https://doi.org/10.1007/s42438-020-00121-8>
- Rudolph, J., Tan, S., & Aspland, T. (2023). Personal digital assistant or job killer? Generative AI and the teaching profession in higher education. *Journal of Applied Learning and Teaching*, 6(2), 7-16. <https://doi.org/10.37074/jalt.2023.6.2.1>
- Tehan, D., & Cash, M. (2020). *Marketplace for online microcredentials*. Joint Media Release, Ministers' Media Centre, Department of Education, Skills and Employment. <https://ministers.dese.gov.au/tehan/marketplace-online-microcredentials>
- Varadarajan, S., Koh, J. H. L., & Daniel, B. K. (2023). A systematic review of the opportunities and challenges of micro-credentials for multiple stakeholders: Learners, employers, higher education institutions and government. *International Journal of Educational Technology in Higher Education*, 20(1), 13. <https://doi.org/10.1186/s41239-023-00381-x>
- Wheelahan, L., & Moodie, G. (2022). Gig qualifications for the gig economy: Micro-credentials and the 'hungry mile.' *Higher Education*, 83(6), 1279–1295. <https://doi.org/10.1007/s10734-021-00742-3>
- Wilson, B. G., Gasell, C., Ozyer, A., & Scrogan, L. (2016). Adopting digital badges in higher education: Scoping the territory. In D. Ifenthaler, N. Bellin-Mularski, & D. K. Mah (Eds.), *Foundation of digital badges and micro-credentials* (pp. 163-177). Springer International Publishing. [https://doi.org/10.1007/978-3-319-15425-1\\_9](https://doi.org/10.1007/978-3-319-15425-1_9)

Copyright: © 2024. Jeremy Hanshaw. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Organisational resilience in a higher education institution: Maintaining academic continuity, academic rigour and student experience in the face of major disruption (Covid-19 pandemic)

---

Adeela Ahmed Shafi<sup>A</sup>

A

*Professor, MBE, School of Education & Humanities, University of Gloucestershire, UK*

Tristan Middleton<sup>B</sup>

B

*Senior Lecturer, School of Education & Humanities, University of Gloucestershire, UK*

---

### Keywords

---

Academic rigour;  
COVID-19;  
higher education;  
organisational resilience;  
student experience.

### Abstract

---

This paper aims to understand how an institution responds to a major disruption such as the Covid-19 pandemic by focusing in detail on one university in England. The study collected data from a range of levels, including survey data from students and staff as well as recruitment data, degree outcomes and financial impact to explore how academic continuity, academic rigour and student experience can be maintained.

---

### Correspondence

ashafi@glos.ac.uk <sup>A</sup>

---

### Article Info

Received 14 August 2023

Received in revised form 21 November 2023

Accepted 25 November 2023

Available online 9 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.4>

Using a systems-based approach and drawing on an organisational resilience framework, findings demonstrated that the case study university had made a positive adjustment to the pandemic. It managed to maintain academic continuity, rigour, and the student experience. What was less clear were the longer-term impacts and the extent of that resilience as defined in the organisational resilience literature which focuses on adversity as an opportunity to learn and land in an overall better place after adversity rather than return to a 'business as before' place. This is applicable to other universities that made similar adjustments in response to the pandemic. A better understanding of organisational resilience in higher education institutions is important in order to enable them to plan for other such disruptions that are part of a modern, connected and global world.

## Introduction and context

Universities in the UK are acquainted with change and adaptability, largely because the sector is the subject of much policy change and regulation, operating within a neoliberal marketised context (Radice, 2013; Brown, 2015; Andrew, 2023). Post-1992 universities – polytechnics that gained university status in 1992 – are further accustomed to responding to policy, market, and technological changes in a bid to compete with Russell Group universities as well as each other (Boliver, 2015). It means that UK universities are systems, which although complex and dynamic, are familiar with external pressures as ‘business as usual’ (Ahmed et al., 2015). However, how well UK universities are prepared for major disruptions through high-impact, low-probability (Sheffi, 2005) events is not well-researched. The number of pre-Covid studies on Google Scholar on UK universities and major disruption is zero, compared to other countries such as the USA (e.g. Kapucu & Khosa, 2013), Malaysia (e.g. Jaradat et al., 2015), New Zealand (e.g. (Kachali et al., 2012) or Taiwan (e.g. Han et al., 2020) with greater risks of natural disasters or other major infrastructure-type disruptions like war or civil unrest. The UK political system has fewer geopolitical risks than other parts of the world, which in some ways contributes to greater resilience but could also be a risk because some level of adversity helps develop resilience (Blyth & Mallett, 2020). The Covid-19 lockdown restrictions imposed in March 2020 by the UK government meant universities had to respond rapidly to continual government changes to ensure rigour of the educational award (Gamage et al., 2020a), integrity (Gamage et al., 2020b) and quality of educational experiences – in other words, ‘academic continuity’ (SchWeber, 2013) – in addition to ensuring ongoing student enrolment (Ahlburg, 2020). Disruption to teaching and learning can lead to “substantial financial loss, reputation damage, job losses, [and] curriculum limitations” (SchWeber, 2008, p. 38), even in the short term. A deeper understanding of universities responses is important for knowing what can be learnt from such disruptions. Grafton et al. (2021) maintain that innovative strategies with individually supportive staff can be successful for academic continuity during disruptions like Covid-19.

This paper is the third in a trilogy of papers centred on a case study of one university in England and how it responded to the global Covid-19 pandemic to enable it to continue its business in a bid to minimise negative impact. This paper explores the university as an organisation and a system, framed within the organisational resilience literature. The study draws on primary student and staff data as well as secondary university-level data on the measures taken to support academic continuity, rigour and student experience. Conclusions will be drawn as to the extent to which the case study university could be deemed resilient based on a review of the literature and what lessons can be learned for future disruptions. It thereby contributes empirical evidence to the growing theoretical field of organisational resilience.

## Literature review

Fast-moving technology, political instabilities, instant communication and global events all mean that change, flux and uncertainty is an inevitable part of modern living, and the associated risk of major disruptions can have more profound effects. These include natural disasters, war, terrorism or pandemics which although lower in probability, have the potential for high impact (Sheffi, 2005). Organisations as well as individuals need to develop mechanisms to ensure that they not only cope and manage in the face of such adversities and disruption but also learn from them and end up in a better position than before (Sutcliffe et al., 2016). Plenty has been said about human resilience at the individual level in the field of psychology (e.g. Ungar, 2013). However, the literature on resilience has broadened out to encompass community resilience, organisational resilience, educational resilience, urban resilience, to mention a few. This study is situated within the relatively new fields of organisational resilience and educational resilience literature. Educational resilience was defined by Wang et al. (1997, p. 2) as “the heightened likelihood of educational success despite personal vulnerabilities and adversities brought about by environmental conditions and experiences”. These adversities are generally related to external conditions such as socioeconomic disadvantages or other issues outside of the control of the individual learner, such as bereavement or disability. Organisational resilience is contextualised within a systems-based approach (Brown et al., 2017) and defined as:

the emergent property of organisational systems that relates to the inherent and adaptive qualities and capabilities that enable an organisation’s adaptive capacity during turbulent periods. The mechanisms of organisational resilience thereby strive to improve an organisation’s situational awareness, reduce organisational vulnerabilities to systemic risk environments and restore efficacy following the events of a disruption (Burnard & Bhamra, 2011 p. 5587).

Here, the importance of identifying opportunity through disruption as a feature of organisational resilience is emphasised. It is in line with wider considerations of organisational resilience with the ability to learn and develop as a result of adversity as a core feature (Rehak, 2020; Bouaziz & Smaoui Hachicha, 2018; Nkwunonwo & Mafimisebi, 2015) and as part of dyadic relationships within wider systems (Sabatino et al., 2016).

This resonates with the principles of the dynamic interactive model of resilience (DIMoR) (Ahmed Shafi et al., 2020), which is about the emergent and agentic nature of resilience. The DIMoR recognises systems (individuals or organisations) within their own right but that they are themselves located within the wider system of community and society, all of which have dynamic and interactive impacts upon them. At the same time, a system also impacts other systems around them and alters the path of external stimuli/systems to meet or even change their own pathway and seek opportunity in the adversity or disruption (see Figure 1). Key conditions within an organisation enable that system and its component parts to be resilient and emerge as a stronger entity, able to proactively seek opportunities (Bouaziz & Smaoui

Hachicha, 2018) through adversity or disruption. In this way, organisational and educational resilience intersect because the university as an organisation can foster conditions to enable the individuals within it to be resilient despite the adversities they may face.

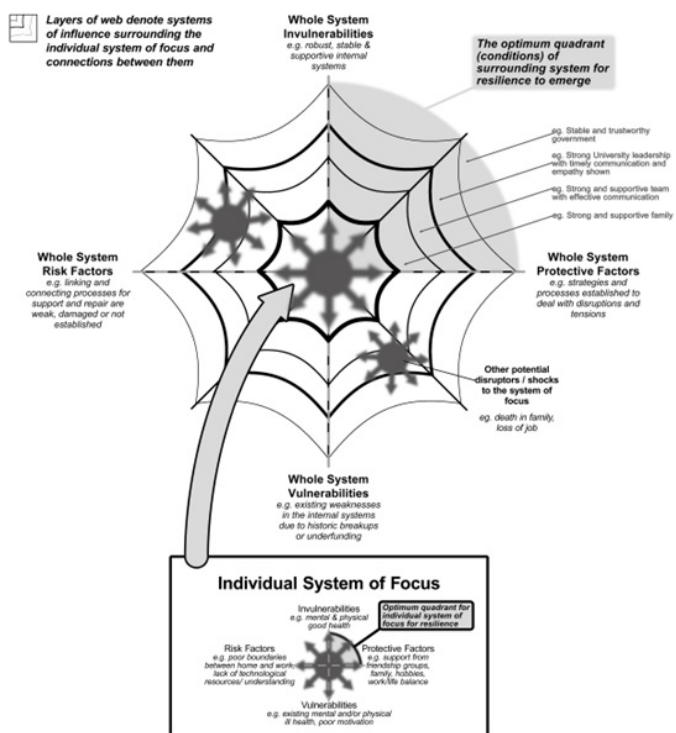


Figure 1. The Dynamic Interactive Model of Resilience (DIMoR) (Ahmed Shafi et al., 2020). The web-like structure illustrates the systems within which an individual may be situated and considers other individual systems within it (represented by the other 'orbs').

The overall model (Figure 1) can represent a university as a system with the orbs being the various actors, such as staff or students. The university would represent the exosystem and meso-system, whereas the web-like structure would represent the systems and processes (e.g. academic regulations, online platforms) within which the various components of the university actors (orbs) operate. The risk-protective aspect of the matrix refers to those elements which can pose as (external) risk or protective factors to the (university) system; these could be the disruptors (e.g. the pandemic) or protective factors (e.g. sound finances). The vulnerabilities-invulnerabilities refer to the (internal) factors such as dwindling student numbers (vulnerabilities) or the impact of the leadership (invulnerabilities). All these factors are inter-connected, interactive and dynamic, shaping the emerging resilience of the system.

Burnard and Bhamra (2011) also present a conceptual model of organisational resilience as a useful theoretical lens. Particularly, they focus on the importance of detection and activation as key features of a resilient organisation, building the ability to have 'positive adjustment' to disruption and adversity (Pratt, 2000), in contrast to previous notions of a more rigid response (Staw et al., 1981) in a bid

to maintain stability. This was later labelled as a 'negative adjustment' (Pratt, 2000) because, ultimately, it restricts the organisation's opportunity for development and can consequently threaten its survival, even having withstood the initial disruption (Chadwick & Raver, 2020).

Nkwunonwo & Mafimisebi (2015) extend the definition of organisational resilience to include transformation of the organisation as a result of adaptive capacity and subsequent application of learning.

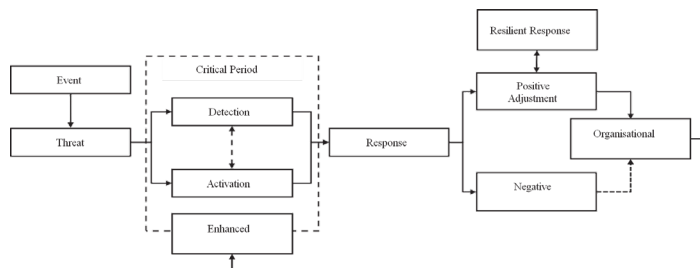


Figure 2. Resilience response framework (Burnard & Bhumra, 2011).

The model (Figure 2) outlines this conceptual model of resilience specifically relating to adverse and disruptive events and Burnard and Bhumra's (2011) critical period, where the response to the event is determined. Here, they emphasise the importance of the detection and activation of the response, which determines whether the organisation will have a negative adjustment (rigid response) or a resilient positive adjustment (flexible approach). Either way, there is opportunity for organisational learning, which should then feed back into the enhanced monitoring of the organisation for detection and activation of future potential similar events.

This framework, combined with DIMoR, gives us a greater understanding of the way in which an organisation can not only capitalise on its inherent features but also respond to adversity and disruption proactively to shape its own resilience. Also useful is the recent work by Dohaney et al. (2020), which specifically explored the characteristics of a resilient university from the perspective of academics. Academics identified the benefits, barriers, and incentives to building resilience in a university at three distinct levels (individual, school, institution), offering insights into resilience-building strategies, which could help move an organisation from possible negative adjustment to positive adjustment (resilience).

These characteristics of resilient academics and institutions contained major themes of communication, community, support, strategic planning, preparedness and leadership that ran across their data. To conclude, Dohaney et al. (2020) argue that leadership is able to create and foster the conditions needed for a resilient organisation, echoing the emphasis placed by DIMoR on how educational settings can play a pivotal role in developing resilience in their learners by being a resilient system themselves.

This current study focused on one university as a case study to explore how as a system it responded to the disruption of the Covid-19 pandemic. We focused on the key elements of

Table 1. Characteristics of resilient academic and resilient institutions (Dohaney et al., 2020).

Resilient academics	Resilient institutions
Are flexible, adaptable, emotionally-resilient, collaborative, empathetic, open-minded individuals (attributes)	Effective communication channels
Respond quickly during a disruption, are digitally literate, organised, prepared and creative-thinkers (capabilities)	A coherent crisis communication strategy
Have a sound awareness of their courses, learner-centred approaches, Learning and Teaching delivery options during disruptions, emergency protocols, and the wider institutional system (knowledge)	An established, coherent, Learning and Teaching disruption plan across all levels of the institution
	<b>Strong resilience-building leadership</b>
	Existing emergency response plans and management
	Existing flexible, blended, and digital learning strategies
	Support for staff to undertake resilience-building initiatives
	Support for staff to develop digital literacy with an effective and easy-to-use digital infrastructure
	A strong sense of staff and learner community

the university system namely, the students, the staff and the systems and processes put in place by university leadership to respond to the impact of the pandemic. The overarching research question was: *How resilient has university been in the wake of the Covid-19 pandemic?* The sub-questions were (i) how has a case study university responded to the disruptor event (Covid-19 pandemic) to maintain academic continuity, rigour and student experience? (ii) How do university staff perceive the university to have responded to the pandemic in terms of student and staff support to maintain academic continuity and student experience? (iii) How do students in a case study university feel it has maintained academic continuity and student experience?

## Methodology

A single case-study design of a post-1992 university in England, with a student cohort of c. 7,950 and staff body of c. 1,500 was employed. The case-study approach enables the blending of description, analysis and the understanding of perceptions (Hitchcock & Hughes, 1995, p. 376), through in-depth analysis and high internal validity (Gagnon, 2010, p. 2). This approach, which focuses on a single unit of analysis (Hammond & Wellington, 2021, p. 20), with an emphasis on context, enables specific contextual factors to be unpicked (Grix, 2018, p. 39) and readers to understand how these ideas fit together (Yin, 2009, pp. 72-73). A pragmatic approach (Biesta, 2020) was adopted to decide what data would best inform understanding without being tied to philosophical dualisms, enabling us to use qualitative and quantitative primary and secondary data.

Primary data were collected from staff and students as part of the data collection for the trilogy of papers on this topic. The development and structure of these surveys are outlined in the first two papers in this trilogy (Ahmed Shafi et al., 2023; Millican et al., 2023). Items that were specifically related to the university responses were extracted, while the full surveys were analysed in the other papers. The student survey had 434 responses across undergraduate and postgraduate cohorts, whilst the staff survey provided 159 responses across academic and other staff. This primary data provided the opportunity to assess, from the perspectives of

students and staff, how successful the university's response was.

The (publicly available) secondary data was supplied by the university but are also available on the UK Higher Education Statistics Agency (HESA), and identified the changes implemented by the university in response to Covid-19 restrictions, referring to retention, progression and achievement rates. The hypothesis was that if student outcomes, student experiences and university finances remained relatively stable despite the pandemic, then it would be reasonable to conclude that the university had made a positive adjustment to the adversity/disruption. Ethical approval for this research was provided by the researchers' university Research Ethics Panel (approval code EDU20209).

## University responses to the restrictions (data point 1)

Like many other universities across the country, the case study university rapidly introduced a number of measures to maintain academic continuity for students whilst seeking to uphold academic rigour and quality of student experience. These included the setting up of a Covid-19 Response Working Group of senior university personnel; moving all teaching and learning online; developing alternative assessments; a new no-detriment policy; relaxing rules for extensions, and a revision of the academic calendar. A number of additional measures designed to support students were also introduced and included support for those in halls during lockdown, early release from contracts, food and provisions support, support for international students to either return home or stay, moving all student support services online, setting up asymptomatic testing centres, as well as increased chaplaincy service and opportunities for people to remain in touch with one another during lockdowns. Each of these measures required a range of system and process adjustments to enable them to happen. The measures described below are included because of their direct impact on academic continuity, rigour and student experience.

### *Covid-19 Response Working Group*

This group drew membership from senior University colleagues and met weekly. However, it sat outside the university's existing Major Incident Plan and the reasoning for this was because of the fast-moving situation of the pandemic.

### *Moving all teaching and learning online*

Moving all teaching and learning online required the relevant IT systems and procedures to adjust to accommodate this. It required rapid training and development for all academic staff to be able to teach and interact with students online, as well as rapid (IT) infrastructure development to support these changes.

### ***Alternative assessments***

This involved the creation of a temporary framework within which module tutors could gain approval for an alternative assessment where the existing assessment type was anything other than coursework (individual) submitted online. To maintain rigour Professional and Statutory Regulatory Bodies were consulted and external examiners were integrated into a newly-designed in-semester assessment scrutiny process.

### ***Extensions***

The extension of the self-certified period of an extension to assessment deadlines from 7 days to 14 days was introduced in order to mitigate the detrimental effects of the ongoing Covid-19 pandemic and provide additional support to students. It also relieved pressure on local doctors/professionals from having to provide evidence to support extensions.

### ***A no-detriment policy***

The development of a no-detriment policy included variations to the Academic Regulations for Taught Provision, whereby counting only the best credit would feature in award classification calculations. This was so that students' final degree classification had no detriment due to the pandemic.

### ***The revision of the academic assessment calendar***

The assessment calendar was revised to accommodate the delays caused by the submission of assessments, which resulted in exam board delays with a potential impact on progression and awards.

### ***Support for staff***

An Agile Working framework was put into place to support staff working through the pandemic. Additional training, guidance and communication were provided for staff in order to implement the changes.

### ***Survey data (data point 2)***

#### ***Staff survey***

This survey was conducted online across academic and professional staff and was designed to understand the impact of the pandemic on staff and how they had coped with the restrictions and the swift changes they had to make in order to support teaching and learning. Items relevant to how staff felt with regards to how the university responded to the pandemic were extracted from the overall survey, which was responded to by 159 members of staff.

### ***Student survey***

The student survey was designed to understand how students were coping with the impact of the pandemic on their lives and how they got on with their studies during this time. The survey was administered online, and participants were reached via the university homepage as well as through their course leaders and other staff. Items relevant to how students felt with regard to how the university responded and supported them through the pandemic were extracted from the full survey. A total of 436 students responded.

### ***Wider organisational impacts (data point 3)***

#### ***Financial***

In addition to data from the university responses to the pandemic, the staff and the student survey data, it was important to assess the impact of the pandemic on the university's finances. The hypothesis here is that if the university's finances had the ability to cope with the additional costs associated with the pandemic with no direct long-term impact, then the university could be considered to have made positive adaptations to the implications of the pandemic.

#### ***Recruitment, retention and awards***

University finances are also connected to the impact on student recruitment, retention and awards and so this too was examined. The hypothesis was that if the university managed to maintain student recruitment, retain students and enable students to exit with awards that did not negatively reflect the impact of the pandemic, then it could be argued that the organisation had made positive adaptations to continue its core business through the measures it took.

#### ***Analysis of data***

To understand the university responses to the impact of the pandemic, they were categorised into whether they were designed to maintain academic continuity, academic rigour or student experience (Table 2).

These responses were then analysed in two stages (Table 3). At Stage 1, the university's responses to the restrictions (data point 1) were compared against Dohaney et al.'s (2020) Characteristics of Resilient Institutions to explore areas where they mapped onto these characteristics and identify any gaps (research question (i)).

For Stage 2, the staff and student data (data point 2) were used to assess the impact of the changes on academic continuity and student experience (research questions (ii) and (iii)). Stage 3 assessed the wider impacts (data point 3) on student outcomes to assess academic rigour and the extent of positive adjustment (overarching research question).

Table 2. The case study university responses were categorised into academic continuity, academic rigour or student experience.

Academic Continuity	Academic Rigour	Student Experience
All teaching and learning moved online.	No detriment policy introduced	Self-certified extensions to assessment deadlines extended from 7 to 14 days
COVID-19 Deferral Process introduced to allow those with extreme circumstances to defer their studies mid-semester	Alternative assessments to retain learning outcomes	Maintenance of physical support/visits for halls students (wellbeing / mental health / practical issues)
Support with IT equipment for academic staff	Revision of academic assessment calendar (to accommodate delays caused by extensions, impacting on exam boards)	Delivery of facilities and maintenance solutions for halls' students during lockdowns
Guidance and communication to academic and professional staff on changes	Additional subscriptions to electronic resources	Immediate switch to virtual appointments so that support for students could continue (Counselling, Disability, Employability, Helpzones, Mental Health, Money Advice, Student Achievement, Wellbeing)
Agile working framework for staff		Range of Chaplaincy events with a variety of virtual and in-person across all sites
Covid-19 Working Group		Setting up asymptomatic testing centres for all students to enable them to return home during the government 'travel window'
		Specific support for international students
		Early release from halls' contracts
		Library and IT changes to support remote studying and access to materials

## The data

This section presents the findings of the three data points relevant to the impact of the changes introduced to ensure academic continuity, rigour and student experience. Data from the university that was available on the number of extensions under the new CV19 scheme, interruptions or withdrawals from the study formed data point 1. The staff and student data that was extracted from the surveys, relevant to how staff or students perceived the changes made by the university, was data point 2. Degree outcomes data, types of exit awards, recruitment data and financial data as data point 3 were examined to assess the impact Covid-19 may have had on these areas.

### Data Point 1: University responses to Covid-19 restrictions

#### Use of extensions

The Extenuating Circumstance 1 (EC1) is a self-certified 7-day extension self-applied by the student via the student account with no external evidence required. The EC2 is a university-applied extension which requires independent verifiable evidence to support an extension of up to four weeks and more in exceptional circumstances. The CV19 replaced the EC1, extending self-certification from 7 days to 14 days during the pandemic. The WA3 is a well-being-based extension and approved by the University with independent verifiable evidence for a time ascertained by senior tutors and other relevant staff at the University based on the student's needs and welfare.

Table 3. Use of extensions.

Year	EC1	EC2	CV19	WA3
2019/20	6755	2163	13252	34895
2020/21	452	6699	26343	40168
2021/22	9584	2994	0	15135

Table 3 shows that extensions were well used by students to support them during Covid-19. The EC1 dropped considerably in 2020/21 because they were replaced by the CV19. In 2021/22 the CV19 was reverted to the EC1.

#### Interruption of studies/withdrawals

The number of interruptions of studies increased during the pandemic but began to come down in 2021/22, suggesting that these interruptions were due to the pandemic and began to reduce as the pandemic eased.

Table 4. Number of interruptions.

Year		UK students	Home	International students
2018/19	Interruption of Studies	33		1
2019/20	Interruption of Studies	180		17
2020/21	Interruption of Studies	197		13
2021/22	Interruption of Studies	168		12

#### Exit awards

Table 5 lists the number of exits per year that left with the intended award, lesser award or no award.

Table 5. Exit awards data.

Year	All Exits	Intended Award	Lesser Award	Exit Award	No Award	% Exit Award	No Award
2015/16	3191	2291	333	567		18%	
2016/17	3558	2549	479	530		15%	
2017/18	3739	2723	465	551		15%	
2018/19	3626	2736	420	470		13%	
2019/20	3620	2729	394	497		14%	
2020/21	3522	2659	304	559		16%	

The data shows that on the whole, the exit award status of students has not changed significantly during the pandemic. It suggests that impacts of the pandemic on students were managed and supported.

### Data Point 2: Survey data

#### Student survey data

Table 6 shows responses from relevant questions in the survey to understand what had helped students in 'maintaining academic continuity'.

Table 6. Student survey on Covid 19 and their studies.

Student survey	
<i>Thinking more specifically about University, who or what has helped you to get on with your studies during the pandemic?</i>	
<b>Having lectures online</b>	236 (54.1%) stated that this helped a lot or a bit although there were a number of students (166) (36.3%) say it did not help much or did not help at all.
<b>The lecturers</b>	324 (74%) students saying they helped a lot or a helped a bit.
<b>Being able to get extensions</b>	266 (61.0%) believe this is helpful though a number of students felt this to be non-applicable (98) (22.5%), which is quite a large proportion and could be due to the nature of their course.
<b>Having some face-to-face teaching where possible</b>	309 (70.8%) found this to be helpful though for a small proportion (40) (9.2%) it was not applicable. It is possible some courses may not have offered any face-to-face teaching in the period before the survey.
<b>My personal tutor</b>	287 (66.1%) found this to be helpful, though for 139 (32.1%) students, they made no difference or did not help at all.
<b>University IT services</b>	For 148 (34.8%), these made no difference, though for 135 (31.2%) they were helpful. For 77 (17.8%) students, they did not help much or at all.
<b>University welfare services</b>	156 (35.9%) said they made no difference and 113 (26%) did not think the service was applicable.
<i>As a result of the pandemic, what things have made studying harder or easier?</i>	
<b>Online learning</b>	290 (66.7%) students found online learning made studying a bit harder or harder which, when taken with Question 4.2 suggests that while the availability of online learning was good, it was hard for students.
<b>Informal contact with lecturers</b>	220 (50.5%) found it harder, though for 94 (21.6%) it made no difference. For 99 (22.8%) it was easier, perhaps because they felt the lecturers were online a lot more, and therefore perhaps more accessible.
<b>Access to resources e.g. library, studios, labs, IT etc</b>	128 (27.2%) students in the sample indicated that access to university resources made no difference to them in terms of studying. 149 (34.4%) said limited access made studying harder.
<i>When things did not go very well, what did you do? Students had to tick all that applied.</i>	
<b>Contact my lecturer</b>	209 (49.9%)
<b>Asky my personal tutor</b>	174 (41.5%)
<b>Ask other students</b>	289 (69%)
<b>Ask family</b>	142 (33.9%)
<b>Contact university welfare</b>	33 (7.9%)
<b>Wait till it happens again</b>	30 (7.2%)
<b>Ignore it</b>	107 (25.5%)
<b>Get anxious</b>	229 (54.7%)
<b>Feel low</b>	173 (41.3%)
<b>Get angry</b>	87 (20.8%)
<b>Do nothing</b>	69 (16.5%)

These data show that having lectures online, contact and support from lecturers and personal tutors were helpful. However, University welfare or IT services were not considered significant. Many found online learning and contacting lecturers harder, although a number did find it easier to contact staff. Further, while many students felt anxious, low or even angry, many consulted other students, housemates, lecturers or personal tutors. It was also clear that some students did not do very much when things did not go well, such as ignoring it or waiting to see if it happened again. These data suggest that overall the pandemic made studying much harder and there are a number of (vulnerable) students who have not done much to seek support despite feeling anxious. Still, where students had a relationship with lecturers or tutors, they felt supported.

Staff largely felt supported by colleagues to carry out their work during the pandemic, though this was patchier when it came to equipment and resources. They also felt less supported with regard to balancing screen time or overall work-life balance (see Millican et al., 2023). Staff did indicate, however, that the University communicated well with staff and with students, with good leadership visibility, provided good IT support as well as positively supporting student well-being. Staff felt students would have found the use of extensions most helpful in supporting their learning.

## Staff survey data

Table 7. Staff survey responses.

Staff Survey	
<i>How did you find the IT or equipment support from the university?</i>	
<b>IT support or equipment from the University</b>	17 (10.8%) found this very challenging though 43 (27.4%) found it a bit difficult. 37 (23.6%) found it manageable and 30 (19.1%) said it made a positive impact (30).
<i>Which of the following helped you to manage the change of moving your work and/or teaching and assessments online? Tick all that apply.</i>	
<b>Course team/colleagues</b>	107 (68.25)
<b>IT support</b>	62 (29.5%)
<b>Online platform guides</b>	65 (41.4%)
<b>Manager</b>	19 (12.1%)
<b>Professional services staff</b>	82 (52.2%)
<b>Working it out myself</b>	121 (77.1%)
<i>How well do you think the University supported you in terms of the following?</i>	
<b>Using the technology</b>	110 (69.3%) felt supported.
<b>Health and well-being</b>	86 (54.1%) felt it to be well supported, though 59 (37.1%) did not.
<b>Working from home</b>	92 (56.3%) felt supported whereas 33 (20.9%) felt somewhat supported.
<b>Learning new ways of doing things</b>	96 (60.4%) felt supported and 33 (20.8%) felt somewhat supported.
<b>Screen time balance</b>	76 (47.8%) did not feel well supported.
<b>Work life balance</b>	60 (37.3%) did not feel this was well supported, though 50 (31.4%) felt it was ok.
<i>How well do you think the University did in the following areas overall? 1= Not very well and 5= Very well</i>	
<b>Communication with staff</b>	111 (69.8%) responded 4 and 5.
<b>Communication with students</b>	107 (67.7%) responded 4 and 5.
<b>IT support</b>	109 (62.7%) responded 4 and 5.
<b>Supporting student well-being</b>	95 (59.7%) responded 4 and 5.
<b>Adopting systems</b>	106 (66.7%) responded 4 and 5.
<b>Senior leadership visibility</b>	78 (49%) responded 4 and 5 with 40 (25.2%) at 1 and 2.
<i>What do you think students will have found the most helpful in terms of supporting their learning during the pandemic? 1= Least helpful and 5= Most helpful</i>	
<b>Extensions, deferral policy, interruption of studies</b>	109 (79.9%) responded 4 and 5.

## Stage 1 and 2 analysis

For stage 1, university responses (data point 1) were mapped against Dohaney's (2020) characteristics of a resilient institution (columns A and B in Table 8 below). For stage 2, data extracted from the student and staff surveys (data point 2) (column C) were used in order to ascertain if there was an indication that the university had that resilient characteristic.

Table 8 identifies that the university demonstrated a number of resilient characteristics as illustrated by the measures introduced to maintain academic continuity, rigour and student experience. The data in Columns C shows that, on the whole, students and staff acknowledged and felt supported by a number of these initiatives. It should be noted that on a few of the characteristics, no specific data were collected.

## Data point 3: Wider impacts

This data point sought to assess the impact of the pandemic on the university's finances. The data was extracted from the HESA data at the University level. Degree outcomes were obtained from the university as the latest data on that would not yet have been published at the time of writing. Recruitment data was obtained from the university and is also publicly available.



## Financial data

The tables below list the key financial indicators of the university taken from the HESA website.

Table 8. University responses (Column B, data point 1) mapped against Dohaney et al.'s (2020) characteristics (Column A) of resilient institutions and assessed using staff and student data (Column C, data point 2).

A. Characteristic of Resilient institutions (Dohaney et al., 2020)	B. Data Analysis Stage 1: University Responses mapped against A.	C. Data Analysis Stage 2: Staff Responses	C. Data Analysis Stage 2: Student Responses	D. Staff and student survey questions
a) Effective communication channels	Guidance and communication to academic and professional staff on changes. (Academic Continuity)	Positive view of communication with staff – 69% Positive view of communication with students - 67%	A small number of negative comments from the qualitative data	Q12.1 Q12.2
b) A coherent crisis communication strategy	Covid-19 Steering Group Major Incident Plan	There was no data on this in the surveys.	There was no data on this in the surveys.	
c) An established, coherent, Learning and Teaching disruption plan across all levels of the institution	All teaching and learning moved online. (Academic Continuity)	Students: 54% stated online learning helped to continue their studies. Staff: 35% perceived staff received positive support for teaching online.	Very small number of positive comments from the qualitative data about the No Detriment policy Just one comment from staff about this area.	Q4.2 Q11.6
	No detriment policy introduced (Academic Rigour)	One qualitative comment about this area.	Changes to assessment formats and the calendar were not commented on.	Q4.4 Q12.6 Q13.7
	Alternative assessments to retain learning outcomes. Revision of academic assessment calendar (Academic Rigour)	Changes to assessment formats and the calendar were not commented on.	Changes to assessment formats and the calendar were not commented on.	
	Self-certified extensions to assessment deadlines extended from 7 to 14 days (Student Experience)	68% saw university changes to extension policies had been handled well and 70% stated these changes had been helpful to students.	61% found ability to get extensions helpful.	
d) Strong resilience-building leadership	Maintenance of physical support/visits for halls' students (Student Experience)	Staff did not make comments about the support provided for students in halls.	Students did not make comments about the support provided for students in halls.	
	Delivery of facilities and maintenance for halls' students during lockdowns Immediate switch to virtual appointments so that support for students could continue	55% of staff had a positive perception of how the provision of support services had been handled and 49% thought that the maintenance of this provision had been positively handled.	17% students perceived that access to university support services had made studying easier during the pandemic, with 38% stating that this opportunity made no difference. 19% stated that university welfare services had been of help. 8% identified university welfare services as a source of support when things didn't go well.	Q12.5 Q13.6 Q5.10 Q4.13 Q6b

	Range of Chaplaincy events with a variety of virtual and in-person across all sites	The chaplaincy was mentioned by 1 member of staff as a positive source of support in the qualitative survey data.	The chaplaincy was mentioned by 1 member of staff as a positive source of support in the qualitative survey data.	
	Setting up asymptomatic testing centres for all students to enable them to return home during the government 'travel window' Early release from halls contracts	This aspect of the university response received no mention from staff.	This aspect of the university response received no mention from students.	
		This aspect of the university response received no comment from staff.	This aspect of the university response received no comment from students.	
e) Existing emergency response plans and management		There was no survey data on this.	There was no survey data on this.	
f) Existing flexible, blended, and digital learning strategies	All teaching and learning moved online. (Academic Continuity)	See section C.	45% of students identified that access to library and other resources had been harder. Within the qualitative responses, continuing library access was identified a number of times.	Q5.8
	Additional subscriptions to electronic resources (Academic Rigour)	Specific changes to library services were not directly commented on by staff.	60% of staff identified that the university positively supported them to learn new ways of doing things, whilst 36% stated that they were positively supported to teach online.	Q11.4 Q11.6 Q4.7 Q4.9
g) Support for staff to undertake resilience-building initiatives		Staff: 54% identified that their own health and well-being was positively supported by the university. 48% identified that the university did not do well to support screen-time balance and 38% that work/life balance was not well supported.	Students: 71% identified the opportunity of having some face-to-face teaching as supporting them with their studies and 66% identified their Personal Tutor as a support.	Q11.2 Q11.2 Q11.13
	Support with IT equipment for academic staff (Academic Continuity)	Staff: 19% identified that IT equipment support made a positive impact, whilst 38% identified that this as a negative.		Q7.7
i) A strong sense of staff and learner community		Staff: 43% identified positive support from the university to deal with student difficulties.	Students: 75% identified individual lecturers as a help to get on with their studies. 69% looked to other students or housemates when things were not going well, in comparison to 50% who would contact a lecturer and 7.9% who would go to	Q11.11 Q4.3 Q6.b

			university welfare services.	
j) University responses without the Dohaney et al (2020) model	Agile working framework for staff (Academic Continuity)	Staff did not identify this as an issue.	Students did not identify this as an issue.	

Table 9. Financial data from the case study university.

	2015/16	1016/17	2017/18	2018/19	2019/20	2020/22
Tuition fee income	5,245	5,828	3,479	2,063	-1,404	967
Total income and education contracts	55,842	60,931	64,791	62,530	60,550	63,971
Total income	76,787	83,000	84,967	83,764	77,834	80,525
Total expenditure	76,010	79,042	82,156	81,086	79,684	80,333
Net cash inflow from operating activities as a % sum of income	8.37	16.18	8.85	7.67	10.95	11.83
Surplus/deficit (as a % of total income)	0.88	4.77	3.31	3.2	-2.38	1.1
Staff costs as a % of total income	55.15	53.82	56.79	58.2	61.97	58.8
Net liquidity days	119.07	130.31	124.17	109.28	128.65	138

This data show that whilst there is a small level of negative variation in the financial position of the university, given the necessary adjustments as a result of the pandemic, the financial footing of the university has remained relatively steady.

### Degree outcomes data

The case study university, like other universities, responded at pace showing support for students and demonstrating an understanding of their anxieties. The temporary 'no detriment' version of the case study university's standard approach to degree classification was approved by the Academic Board of the university. The intention was to ensure that the degree classifications of students graduating at the end of 2019/20 and in 2020/21 were not adversely affected by lower results in assessments completed during lockdown. Table 10 shows the degree classifications awarded at the university in 2017/18 and 2018/19 according to the rules in force for those years (the 'legacy algorithm') and the classifications which would have been awarded in 2019/20 had the planned algorithm been used (the '2017/18 algorithm').

Table 10. Degree outcomes had there been no 'no-detriment' policy.

Class	2017/18		2018/19		2019/20	
1	465	26.9%	465	27.2%	284	18.8%
2:1	881	51.0%	858	50.2%	756	50.1%
2:2	353	20.4%	345	20.2%	423	28.1%
3	28	1.6%	41	2.4%	45	3.0%
TOTALS	1,727	100.0%	1,710	100.0%	1,508	100.0%

Table 10 shows that degree classifications would have been negatively affected if the original planned algorithm had been used for 2019/20. Table 11 shows the actual degree classifications as awarded with the 'no-detriment' policy for 2019/20 and the legacy algorithm for 2017/18 and 2018/19.

Table 11. Degree outcomes with the 'no-detriment' policy for 2019/20.

Class	2017/18		2018/19		2019/20	
1	465	26.9%	465	27.2%	547	36.3%
2:1	881	51.0%	858	50.2%	741	49.1%
2:2	353	20.4%	345	20.2%	214	14.2%
3	28	1.6%	41	2.4%	6	0.4%
TOTALS	1,727	100.0%	1,710	100.0%	1,508	100.0%

The figures from Tables 10 and 11 reveal that there was in fact a significant increase in the number of upper degrees awarded in 2019/20 compared to previous years. The actual increase was from 77.4% in 2018/19 to 85.4% in 2019/20, an increase of 8.0 percentage points. So, in terms of achieving its aim, the 'no-detriment' policy had worked to a greater extent than with the 'legacy algorithm'.

### Recruitment data

Table 12 shows that applications, offers, conversions and enrolment remained stable throughout the pandemic at the case study university. This is in spite of the landscape of higher A-Level grades following the use of teacher assessment and the consequent increased entry of students to Russell Group/red-brick universities.

Table 12. Undergraduate recruitment data from the case study university.

Year		Applications	Unconditional Firm Offers	Enrolments	Applications to enrolment	UF to enrolment
2017/18	UG	9900	2672	2620	26%	98%
2018/19	UG	8937	2437	2368	26%	97%
2019/20	UG	8266	2324	2245	27%	97%
2020/21	UG	9487	2721	2621	28%	96%
2021/22	UG	9596	2479	2353	25%	95%

### Stage 3 analysis

This stage of analysis explored the wider university data (data point 3) to understand the extent to which academic rigour and student experience were maintained through the pandemic. In summary, the data show that whilst the university did experience challenges during the pandemic, the measures introduced and how the university responded meant that it did not suffer in terms of financial difficulty, degree outcomes, or overall undergraduate recruitment.

### Discussion

The findings from this study demonstrate that the case study university's responses to the pandemic to maintain academic continuity (SchWeber, 2013), rigour (Gamage et al., 2020a), and student experience were largely successful when explored through the Resilient Organisations Framework of Dohaney et al. (2020) and the Burnard and Bhumra (2011) model. This was evidenced through wider impacts such as the financial position of the university, undergraduate

recruitment data and degree outcomes, all of which could be considered indicators of university success and 'academic continuity' (SchWeber, 2008, 2013). It suggests that the university had made a positive adjustment (Pratt, 2000) to its usual business in the face of adversities caused by the pandemic.

Connecting this to the DIMoR model (Ahmed Shafi et al., 2020), it could be argued that the case study university, as the overall system, effectively managed the various components of its system. This included students and staff as key actors within it and the impact of the pandemic on them both as individuals of part of the system and also in recognition of the wider systems (family and community) of which they were a part. The data from the Staff and Student surveys support the view that the university's approach fostered a resilience-promoting environment in the context of the pandemic. Further, wider measures such as finances, degree outcomes and recruitment data also support this. The range of adaptations made by this and many other UK universities reflect the acceptance of the inter-connected, interactive and dynamic nature of systems (Ahmed Shafi et al., 2020) where innovative strategies can contribute to academic continuity (Grafton et al., 2021). Such a 'systems-based' approach could explain why the university made a 'positive adjustment' (Pratt, 2000) and an emergent resilience indicated by the findings presented in data points 2 and 3 and assessed using Dohaney et al.'s (2020) framework.

Drawing on the Burnard and Bhumra (2011) model (Figure 2), the findings from this study show that the university did demonstrate a positive adjustment and a resilient response. However, it should be noted that the organisational learning element that then feeds into the enhanced monitoring stage of the framework was not evident, except perhaps in the Agile Working Framework, which was adopted even after the pandemic. All other aspects of the Covid-19 specific responses have reverted, including the CV19 extension, No-Detriment Policy, Interruption of Studies, and face-to-face teaching and learning has resumed, even after a period of hybrid teaching and learning. Perhaps this is because some elements of the temporary changes are not supported by the systems and structures for it to be an ongoing feature. Interestingly, this was also reflected in how the specially formed Covid-19 Working Group sat outside the university's Major Incident Plan, thereby positioning the pandemic as outside the existing plans to address unexpected issues. In this way, it could be argued that the university may be displaying what Straw et al. (1981) described as a rigid response in that the idea is to return to 'normal' or a 'negative adjustment' (Pratt, 2000) after the adversity rather than necessarily moving forward into a new and improved place by 'detecting and scanning' for opportunities the adversity presents.

The current data do not show whether the university demonstrated resilience in the way argued as being key in the recent resilient literature with regard to positive adjustment where an organisation not only survives the adversity but also seeks opportunities for doing things better as a result of that adversity (Burnard & Bhumra, 2011; Sutcliffe et al., 2016; Ahmed Shafi et al., 2020). Firstly, this data is difficult to ascertain because it is perhaps still too

soon after the height of the pandemic. Secondly, it would require a different level and measure of analysis (Kapuca & Khosa, 2013). Thus, though a conclusion can be drawn of positive adjustment, whether this materialises into resilience where the transformation of the organisation occurs as a result of adaptive capacity and subsequent application of learning (Nkwunonwo & Mafimisebi, 2015) is not possible to ascertain with the currently available data.

Disease with the extent of impact as seen in the recent Covid-19 pandemic is a high-impact, low-probability event (Sheffi, 2005), particularly for a country like the UK, and so it could be argued that it would be less prepared. Sutcliffe et al. (2015) posit organisations need to develop mechanisms to ensure that they not only cope and manage in the face of disruption but also learn from them, where they end up in a better position than before the disruption. Identifying opportunity through disruption is an important feature of organisational resilience, and the literature (Rehak, 2020; Bouaziz & Smaoui Hachicha, 2018; Nkwunonwo & Mafimisebi, 2015) points to this as a core feature of a dyadic relationship with wider systems (Sabatino et al., 2016). Whilst this is not necessarily evident from the available data, what is available is that the university, whilst retaining some of the practices from the adjustments made for the pandemic, is more focused on returning largely to the position it was in before the pandemic. This suggests that universities (at least the case study university) in the UK could consider how they use this opportunity for growth and innovative ways for teaching and learning, which could indeed change the face of higher education and address, for example, issues of wider participation and the government levelling up agenda or other ways which could improve the access and quality of higher education.

## Conclusion

It could be argued that one of the reasons that a university was able to operate in the Covid-19 conditions was because (institutional and physical) infrastructures generally remained intact. Universities in higher-risk geopolitical locations (e.g. Fillmore et al., 2011; Kachali et al., 2012; Han et al., 2020) are more likely to be overall resilient than a UK university which enjoys relative political and social stability (Blythe & Mallett, 2020). Perhaps this has enabled the resilience (or positive adjustment). What is less clear are the longer-term impacts with regard to resilience as an opportunity to learn and land in an overall better place after adversity rather than return to a 'business as before' place.

In light of these findings, it is suggested that universities would benefit from not only analysing their own immediate responses to major disruption and the subsequent outcomes in terms of student achievement, retention and recruitment in the short-term but also considering longer-term resilience. Through using the resilience models employed in this research (Ahmed Shafi et al., 2020; Burnard & Bhamra, 2011; Dohaney et al., 2020) universities can begin to understand whether they have 'weathered the storm' simply to return to their pre-pandemic positions or have learnt from and through the challenges presented to reach a forward-looking position of greater resilience.

Major disruptions should not be seen as a storm to weather but also an opportunity to create better higher education institutions. Integrating and embedding characteristics of agile, resilient organisations will not only ensure longevity and resilience but contribute to developing resilient communities going forwards into the future. A better understanding of organisational resilience in higher education institutions would be an important future endeavour in order to enable them to plan for other such disruptions that are part of a modern, connected and global world.

## Acknowledgements

We would like to acknowledge colleagues Richard Millican, Sian Templeton and Jenny Hill for their contributions in collecting the data which this paper draws upon.

## References

Ahlburg, D. A. (2020). Covid-19 and UK universities. *The Political Quarterly*, 91(3), 649–654. <https://doi.org/10.1111/1467-923X.12867>

Ahmed, J. U., Ahmed, K. U., Shimul, Md. A. S., & Zuñiga, R. (2015). Managing strategies for higher education institutions in the UK: An overview. *Higher Education for the Future*, 2(1), 32–48. <https://doi.org/10.1177/2347631114558189>

Ahmed Shafi, A., Templeton, S., Middleton, T., Millican, R., Vare, P., Pritchard, R., & Hatley, J. (2020). Towards a dynamic interactive model of resilience (DIMoR) for education and learning contexts. *Emotional and Behavioural Difficulties*, 25(2), 183–198.

Ahmed Shafi, A., Middleton, T., Millican, R., Templeton, S., Hill, J., & Jones, C. (2023). Learning in a disrupted environment: Exploring higher education student resilience using the dynamic interactive model of resilience. *Journal of Applied Learning and Teaching*, 6(2), 172–186. <https://doi.org/10.37074/jalt.2023.6.2.18>

Andrew, M. B. (2023). Neo-neoliberalist capitalism, intensification by stealth and campus real estate in the modern university in Aotearoa/New Zealand. *Journal of Applied Learning and Teaching*, 6(2), 393–401. <https://doi.org/10.37074/jalt.2023.6.2.16>

Biesta, G. (2020). *Educational research: An unorthodox Introduction*. Bloomsbury.

Blyth, M., & Mallett, S. (2020). Epidemics and pandemics: Effects on societal and organisational resilience. *Journal of Business Continuity & Emergency Planning*, 14(1), 17–36.

Boliver, V. (2015). Are there distinctive clusters of higher and lower status universities in the UK? *Oxford Review of Education*, 41(5), 608–627. <https://doi.org/10.1080/03054985.2015.1082905>

Bouaziz, F., & Smaoui Hachicha, Z. (2018). Strategic human resource management practices and organizational

resilience. *Journal of Management Development*, 37(7), 537–551. <https://doi.org/10.1108/JMD-11-2017-0358>

Brown, N. A., Rovins, J. E., Feldmann-Jensen, S., Orchiston, C., & Johnston, D. (2017). 'Exploring disaster resilience within the hotel sector: A systematic review of literature'. *International Journal of Disaster Risk Reduction*, 22, 362–370.

Brown, R. (2015). The marketisation of higher education: Issues and ironies. *New Vistas*, 1(1), 4–9.

Burnard, K., & Bhamra, R. (2011). Organisational resilience: Development of a conceptual framework for organisational responses. *International Journal of Production Research*, 49(18), 5581–5599.

Chadwick, I. C., & Raver, J. L. (2020). Psychological resilience and its downstream effects for business survival in nascent entrepreneurship. *Entrepreneurship Theory and Practice*, 44(2), 233–255. <https://doi.org/10.1177/1042258718801597>

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th Ed.). Routledge.

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approaches*. Sage.

Dohaney, J., de Róiste, M., Salmon, R. A., & Sutherland, K. (2020). Benefits, barriers, and incentives for improved resilience to disruption in university teaching. *International Journal of Disaster Risk Reduction*, 50, 101691. <https://doi.org/10.1016/j.ijdr.2020.101691>

Duchek, S. (2020). Organizational resilience: A capability-based conceptualization. *Business Research*, 13(1), 215–246. <https://doi.org/10.1007/s40685-019-0085-7>

Engineering. (1996). *Engineering within ecological constraints*. National Academies Press.

Gagnon, Y.-C. (2010). *The case study as research method: A practical handbook*. Québec Que.: Presses de l'Université du Québec.

Gamage, K. A. A., Pradeep, R. G. G. R., Najdanovic-Visak, V., & Gunawardhana, N. (2020a). Academic standards and quality assurance: The impact of COVID-19 on university degree programs. *Sustainability*, 12(23), 10032. <https://doi.org/10.3390/su122310032>

Gamage, K. A., Silva, E. K. de, & Gunawardhana, N. (2020b). Online delivery and assessment during COVID-19: Safeguarding academic integrity. *Education Sciences*, 10(11), 301.

Gibson, C. A., & Tarrant, M. (2010). A 'conceptual models' approach to organisational resilience. *Australian Journal of Emergency Management*, 25(2), 6–12.

Grafton, E. T., Elder, E., & Burton, R. (2021). Innovative strategies to maintain nursing students' academic continuity during the COVID 19 pandemic. *Journal of Applied Learning and Teaching*, 4(1), 21–28. <https://doi.org/10.37074/>

- Grix, J. (2019). *The foundations of research* (3rd ed.). London: Red Globe Press.
- Hammond, M., & Wellington, J. J. (2021). *Research methods: the key concepts* (2nd ed.). Routledge.
- Han, E., Chiou, S.-T., McKee, M., & Legido-Quigley, H. (2020). The resilience of Taiwan's health system to address the COVID-19 pandemic. *EClinicalMedicine*, 24, 100437.
- Hitchcock, G., & Hughes, D. (1995). *Research and the teacher* (2nd ed.). Routledge.
- Jaradat, A., Mziu, H., & Ibrahim, J. (2015). Disaster preparedness in universities. *International Journal of Computer Trends and Technology*, 19, 1–4. <https://doi.org/10.14445/22312803/IJCTT-V19P101>
- Kachali, H., Stevenson, J. R., Whitman, Z., Seville, E., Vargo, J., & Wilson, T. (2012). Organisational resilience and recovery for Canterbury organisations after the 4 September 2010 earthquake. *Australasian Journal of Disaster and Trauma Studies*, 1(1), 11–19.
- Kapucu, N., & Khosa, S. (2013). Disaster resiliency and culture of preparedness for university and college campuses. *Administration & Society*, 45(1), 3–37. <https://doi.org/10.1177/0095399712471626>
- Millican, R., Templeton, S., & Hill, J. L. (2023). Exploring the impact of disruption on university staff resilience using the dynamic interactive model of resilience. *Journal of Applied Learning and Teaching*, 6(1), 1-13. <https://doi.org/10.37074/jalt.2023.6.1.19>
- Nkwunonwo, U., & Mafimisebi, O. (2015). Environmental risk: Exploring organizational resilience and robustness. *International Journal of Scientific and Engineering Research*, 6(1), 1103-1115.
- Radice, H. (2013). How we got here: UK higher education under neoliberalism. *ACME: An International Journal for Critical Geographies*, 12(2), 407–418.
- Rehak, D. (2020). Assessing and strengthening organisational resilience in a critical infrastructure system: Case study of the Slovak Republic. *Safety Science*, 123, 104573. <https://doi.org/10.1016/j.ssci.2019.104573>
- Resilience. (n.d.). Assessing and strengthening organisational resilience in a critical infrastructure system: Case study of the Slovak Republic—*ScienceDirect*. <https://www.sciencedirect.com/science/article/pii/S0925753519321848>
- Sabatino, S., Frangopol, D. M., & Dong, Y. (2016). Life cycle utility-informed maintenance planning based on lifetime functions: Optimum balancing of cost, failure consequences and performance benefit. *Structure and Infrastructure Engineering*, 12(7), 830–847.
- SchWeber, C. (2008). Determined to learn: Accessing education despite life-threatening disasters. *Journal of Asynchronous Learning Networks*, 12(1), 37–43.
- SchWeber, C. (2013). Survival lessons: Academic continuity, business continuity, and technology. In P. Van den Bossche, W. H. Gijssels, & R. G. Milter (Eds.), *Facilitating learning in the 21st century: Leading through technology, diversity and authenticity* (pp. 151–163). Springer Netherlands. [https://doi.org/10.1007/978-94-007-6137-7\\_9](https://doi.org/10.1007/978-94-007-6137-7_9)
- Sheffi, Y. (2005). Preparing for the big one [supply chain management]. *Manufacturing Engineer*, 84(5), 12-15. [https://digital-library.theiet.org/content/journals/10.1049/me\\_20050503](https://digital-library.theiet.org/content/journals/10.1049/me_20050503)
- Sutcliffe, K. M., Vogus, T. J., & Dane, E. (2016). Mindfulness in organizations: A cross-level review. *Annual Review of Organizational Psychology and Organizational Behavior*, 3(1), 55–81. <https://doi.org/10.1146/annurev-orgpsych-041015-062531>
- Ungar, M. (2013). Resilience, trauma, context, and culture. *Trauma, Violence, & Abuse*, 14(3), 255–266.
- Watermeyer, R., Crick, T., Knight, C., & Goodall, J. (2021). COVID-19 and digital disruption in UK universities: Afflictions and affordances of emergency online migration. *Higher Education*, 81(3), 623–641.
- Watermeyer, R., Shankar, K., Crick, T., Knight, C., McGaughey, F., Hardman, J., Suri, V. R., Chung, R., & Phelan, D. (2021). 'Pandemia': A reckoning of UK universities' corporate response to COVID-19 and its academic fallout. *British Journal of Sociology of Education*, 42(5–6), 651–666. <https://doi.org/10.1080/01425692.2021.1937058>
- Weller, M., & Anderson, T. (2013). Digital resilience in higher education. *European Journal of Open, Distance and e-Learning*, 16(1), 53.
- Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Do graduate courses in a HyFlex mode foster emotional, cognitive and behavioral engagement? A consideration

David Santandreu Calonge <sup>A</sup>	A	<i>Department of Academic Development, Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, United Arab Emirates</i>
Mariam Aman Shah <sup>B</sup>	B	<i>Department of Educational Research, Lancaster University, Lancaster, United Kingdom</i>
Melissa Connor <sup>C</sup>	C	<i>UniSA Education Futures, Centre for Research in Educational and Social Inclusion, The University of South Australia, Adelaide, SA, Australia</i>
Patrik Hultberg <sup>D</sup>	D	<i>Department of Economics and Business, Kalamazoo College, United States</i>
Pablo Medina Aguerrebere <sup>E</sup>	E	<i>Department of Communication and Media, Canadian University Dubai, Dubai, United Arab Emirates</i>

### Keywords

Engagement;  
graduate level;  
HyFlex;  
social belonging.

### Abstract

The global pandemic accelerated the adoption of online, blended, and Hybrid Flexible (HyFlex) teaching and learning modalities. The long-term impacts of these changes are still being assessed, but most research to date has focused on undergraduate students in an online setting. In contrast, this study reviews recent research pertaining to HyFlex engagement strategies used by academic staff and personnel with teaching responsibilities at the graduate level and considers how HyFlex courses, the combination of face-to-face instruction and online activities, may foster equivalent learning outcomes, as well as comparable emotional, cognitive, and behavioral engagement. Using a critical reflective approach, the study finds that graduate-level courses taught in a HyFlex mode can offer equivalent learning outcomes, but such outcomes require academic staff development and purposefully designed activities that directly promote emotional, cognitive, and behavioral engagement. Several practical strategies and suggestions to improve engagement are offered.

### Correspondence

[david.santandreu@mbzuai.ac.ae](mailto:david.santandreu@mbzuai.ac.ae) <sup>A</sup>

### Article Info

Received 9 September 2023  
Received in revised form 20 February 2024  
Accepted 21 February 2024  
Available online 1 March 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.18>

## Introduction

Student engagement is a multifaceted construct and a critical factor for academic performance and student retention (Burke, 2019; Pechenkina et al., 2017; Zepke & Leach, 2010). Student *engagement* is often defined as students' willingness and desire to participate and succeed in the learning process. Student *online learning engagement* has been defined by Yang et al. (2018) as "students' devotion of time, energy, value/interest, attitude, learning strategy or even creative thinking in e-learning environments and the motivational and action processes elicited" (p. 3). Research by Thomas et al. (2014) indicated that "staff and students expressed greater satisfaction with online courses that successfully fostered a sense of belonging among students" (p. 76). A sense of belonging influences a person's emotional and cognitive patterns and is considered key to improving academic motivation, (social) integration, satisfaction (Stephens & Morse, 2022), achievement (Pedler et al., 2021) and retention rates (Peacock et al., 2020). Strayhorn (2018) found that graduate students thrived and excelled "where they feel like they belong" (Strayhorn, 2018, p. 138).

As higher education cautiously enters a post-COVID-19 pandemic era, considerable literature has explored student engagement in various online formats such as Emergency Remote Teaching (Calonge et al., 2022), blended (Lim et al., 2007), and Hybrid Flexible (HyFlex) (Bockorny et al., 2023; Raes et al., 2020). HyFlex courses combine face-to-face (F2F) and online activities and remote participation through video capture technology, allowing social distancing, more equitable access, choice, control, and flexibility. Although research has studied the undergraduate student experience with HyFlex, a literature search revealed few studies that specifically focused on academic staff development in the HyFlex modality, especially at the graduate level. Practical synchronous HyFlex engagement strategies fostering emotional, cognitive, and behavioral engagement at the postgraduate level have also received minimal attention in the recent research literature. Considering this, the purpose of this paper is twofold: (1) to review recent research pertaining to HyFlex engagement strategies used by academic staff in tertiary education courses, specifically at the graduate level, and (2) to propose practical suggestions to help improve (a) emotional, cognitive, and behavioral engagement and (b) sense of belonging in graduate HyFlex courses.

## Theoretical framework

The theoretical framework adopted is Fredricks et al.'s (2004) three-dimension engagement framework. Fredricks et al. (2004) introduced a three-dimensional engagement framework that has significant relevance to articles discussing the HyFlex pedagogy. This framework focuses on three critical dimensions of student engagement: emotional, behavioral, and cognitive. Emotional engagement examines students' affective reactions to learning, encompassing their interest, enjoyment, and motivation. Behavioral engagement involves the active participation and involvement of students in educational activities. Cognitive engagement pertains to the extent to which students invest their cognitive efforts in understanding and processing course content. When

applied to the context of articles discussing the HyFlex pedagogy—a flexible approach that combines in-person and online learning—the framework's dimensions become instrumental. HyFlex pedagogy demands a high degree of student autonomy and responsibility in managing their learning experiences, making emotional engagement essential to maintain motivation in both online and in-person settings. Behavioral engagement becomes crucial as students navigate various learning modes, requiring them to actively participate regardless of the format. Lastly, cognitive engagement is at the core of effective learning in HyFlex, as students must deeply process information across different modalities. Therefore, Fredricks et al.'s (2004) engagement framework provides a valuable lens to analyze and enhance student experiences within the multifaceted environment of HyFlex pedagogy. Equivalency Theory, proposed by Simonson et al. (1999), is also reflected in the context of this article through engagement in the HyFlex modality. The theory stipulates that online and face-to-face students will achieve equivalent learning outcomes only when they are offered equivalent/comparable learning experiences (regardless of the method of delivery). That is, to achieve equivalence, "course designers [should] create learning experiences of equivalent value for learners regardless of the course delivery medium, allowing that the experiences themselves could be different" (Lapsley et al., 2008, p. 3).

## Methodology

This study undertakes a critical reflective approach as its methodological base. It examines the global phenomenon of changes in higher education resulting from COVID-19, as documented in the current literature. The authors opted for a scoping review methodology. A broad examination of the published literature was conducted on the topic without strictly predefined criteria for inclusion or exclusion of studies, aiming to provide an overview of the existing literature and identify the breadth and depth of available evidence. The authors engaged with literature to identify themes and gaps in knowledge. Critical reflection is a widely recognized methodology that can provide an evidence-based "framework for deconstructing... assumptions about knowledge" (Hickson, 2015, p.308). Thompson and Thompson (2018) argued that "theorizing practice is at the heart of reflective practice" (p.x). Its use in professional learning and teaching studies, therefore, assists "practitioners to improve practice" (Fook, 2011, p.55) by describing, critically analyzing, synthesizing, evaluating, and devising a plan of action (Gibbs, 1988).

The methodological process of critical reflection was guided by the primary research question:

RQ1. Can graduate courses taught in a HyFlex mode foster equivalent emotional, cognitive, and behavioral engagement as in the F2F modality?

To identify practical strategies for graduate students, one supplementary question was considered:

RQ1-1. Is a sense of belonging relevant to graduate students in HyFlex courses?

The study proceeds as follows: first, guided by the critical reflective methodological approach, relevant literature pertaining to the HyFlex modality, student engagement, and online engagement at the graduate level is reviewed; second, findings are presented; third, reflecting on the supplementary question, practical suggestions to help improve (a) emotional, cognitive, and behavioral engagement and (b) sense of belonging in graduate HyFlex courses are offered. The final section draws conclusions from this study.

## Background

### The HyFlex modality

Although flexible learning modalities have been part of higher education for some time, they have evolved in response to advancements in technologies and environmental changes that demand flexible online learning and teaching options. However, in the past, the need for flexible learning and alternatives to face-to-face teaching were not as significant as they are today. Lockdown conditions advanced the need for and the use of video conferencing technologies, which created a global educational need for better synchronous and asynchronous learning and teaching options, culminating in the *Hybrid-Flexible modality*. HyFlex incorporates elements of blended, online, and hybrid pedagogies by providing students with the flexibility to choose how they engage with peers and professors, course materials and activities, whether in person or online. Particularly, HyFlex extends the flexibility of blended learning by allowing students to choose whether to attend classes in person, participate online, or do both. This flexibility accommodates varying student needs, preferences, and circumstances. In a HyFlex model, students also have the option to participate fully online if they prefer or if circumstances prevent them from attending in person, which is often the case for graduate students. HyFlex is thus a specialized form of hybrid learning, where students have the flexibility to choose the mode of participation for each class session. They can opt to attend in person, participate online, or switch between modes as needed. In short, HyFlex emphasizes student choice and flexibility to a greater extent and requires a more intricate integration of technology to support simultaneous participation from both in-person and online students. Hybrid learning, on the other hand, may involve a predetermined schedule and mode of attendance with less emphasis on student choice and simultaneous engagement. Despite HyFlex being introduced by Beatty (2007) prior to the pandemic, lockdown conditions prompted Kohnke and Moorhouse (2021) to describe HyFlex as a “new” and therefore, a “rarely implemented mode” of learning and teaching (p. 232).

By combining face-to-face and online learning methods, HyFlex has been described in multiple studies as uniquely adaptable to social distancing measures due to its primary characteristics of flexibility and choice for students on how (and where) they engage with a course (Trotter and Qureshi, 2023; Bozan et al., 2023; Detyana et al., 2023; Nelson et al., 2022; Heilporn & Lakha, 2021; Kohnke & Moorhouse, 2021; Romero-Hall & Ripine, 2021; Wilson & Alexander, 2021). Abdelmalak and Parra (2016) argued that “graduate students

as adult learners need flexible instruction that extends the boundaries of learning so that learning can occur in the classroom, in the home, and in the workplace” (p. 23).

HyFlex, according to Beatty (2007, 2019), caters to flexibility in engagement and equivalence in learning through the reusability of learning materials that can be implemented across multiple technology options. Courses incorporating video conferencing technologies (e.g. Zoom) along with in-person instruction, provide students with the ability to select synchronous, asynchronous online or face-to-face options and thus have “greater control over their learning and course engagement modes” (Kohnke & Moorhouse, 2021, p. 232). Given this, HyFlex is viewed as being “learner-centered” as the variety of course engagement modalities shift the design principles away from being instructor-focused, in contrast to more traditional blended learning course designs (Wilson & Alexander, 2021, p. 44).

### Student engagement

To improve quality and learning outcomes, student engagement in higher education has become a priority (Fernández-García et al., 2021). Engagement connects experiences across three key dimensions: emotional, behavioral, and cognitive. As suggested by Subramanian and Mahmoud (2020), tertiary institutions should identify different aspects of student engagement, its main dimensions and problems, the most important factors influencing students’ motivation, and evaluation methods to improve engagement learning outcomes. Previous research suggests that emotional intelligence, resilience, and positive psychology played key roles in this process (García-Martínez et al., 2021). In addition, it is critical that universities support different groups of students, such as international and low-income students. International students may face linguistic and cultural barriers that make it difficult for them to be active learners (Calonge et al., 2023), which suggests a need for initiatives that promote inclusivity, multilingualism, and multiculturalism (Xiao, 2021). All students, but perhaps especially low-income/disadvantaged students may need proactive support mechanisms, academically validating practices in the classroom, as well as customized programs (Swanson et al., 2021).

As academic staff and staff with teaching responsibilities (instructors, teaching assistants, etc.) play a key role in connecting the three dimensions to enhance student engagement, it is important that they adapt to higher education’s new requirements, implement innovative learning and teaching strategies such as Challenge-Based Learning, and consider the role of both emotional and behavioral engagement in the courses they teach (Fernández-García et al., 2021). In addition, student performance and engagement increase when they feel integrated into a community (Durón et al., 2018), when a positive dialogue is cultivated, and where a supportive culture of mental health is promoted with clearly established protocols. To facilitate such integration, instructors can launch inclusive events and workshops that promote well-being and help students develop key transferable skills such as teamwork and or socialization (Martín-Hernández et al., 2021) using



a Community of Inquiry (CoI) format (Garrison et al., 1999). Universities should also develop continuing professional development opportunities that help academics integrate interactive learning strategies (online and F2F) into their practice (Oviawe, 2020).

Tertiary institutions have developed online courses to better satisfy students' and academia's needs (Chen & Yang, 2022; Zou et al., 2021). Students' engagement with online courses is essential and necessitates providing students with Teaching Assistants (TAs) and technical support to avoid frustration and negative experiences (Rajabalee & Santally, 2021). To mitigate the impact of those negative experiences, universities need to analyze how academics' expectations (self-efficacy, academic engagement) and students' needs (knowledge, motivation) integrate into this online environment and propose adequate solutions when they do not match or fit (Lobos Peña et al., 2021).

### Online engagement at the graduate level

Graduate students in online courses respond to teaching and learning that is emotionally connected. Moreover, graduate students are self-determined adult learners and are more likely to have previous (local or overseas) experience of study through undergraduate degrees, as well as professional experience (Calonge et al., 2023; Holzweiss et al., 2014). Graduate students expect a teaching and learning environment that includes research and advanced content and is focused on skill development for a specific professional field (Holzweiss et al., 2014). To engage graduate students online, it is helpful to include a) experiential learning activities, and b) professional and previous experience to devise individual learning goals, as is using teaching approaches of sharing experiences and taking responsibility for facilitating class discussion (Heilporn et al., 2021). Using learning and teaching strategies that draw on graduate students' existent skills and competencies can thus increase emotional engagement in online learning. Teachers of graduate students should use authentic online learning activities such as problem/challenge-based learning and competency tests to increase engagement (Kuchinski-Donnelly & Krouse, 2020). In short, adult learners in graduate courses respond well to online teaching that is (authentic) practitioner-experience-based, passionate and has emotional investment, from both the facilitator and classmates (McDavid & Shepherd, 2021).

### Findings based on literature

In response to the primary research question, "Can graduate courses taught in a HyFlex mode foster equivalent emotional, cognitive, and behavioral engagement as in the F2F modality?", the findings examine the literature and utilize Fredericks et al. (2004) three-dimension engagement framework to discuss a sense of belonging, cognitive load, behavioral engagement, and engagement in the HyFlex modality.

### Emotional engagement and a sense of belonging online

Emotional engagement relates to the student's feelings of interest, boredom, excitement, and anxiety in the classroom (Fredricks et al., 2004). Wang and Huang (2018) found that students felt excluded in blended learning classes as some instructors tended to disregard or failed to remember [online] students when engaging with [face-to-face] students. Similarly, research by Cunningham (2014) indicated that "both groups (online and campus students) reported seeing the other group as quite separate from themselves" (p. 40). In short, online students felt left out or isolated, despite connections with the course content, instructors (interactions/feedback) and peers that could potentially lead to a sense of connectedness and belonging. Conrad et al. (2022) indicated that even though "many instructors offered synchronous lectures or sessions, not all students were able to participate, and while these were recorded live sessions, it never provided the feeling of being socially present in the same space as other students and the instructor." This sense of disconnectedness, [them and us], felt by online students partially contradicts Raes et al. (2020) argument that Hyflex could be a more flexible and more engaging learning space (than fully online or face-to-face). It also casts doubts on *reusability* and *equivalent learning*, two of the guiding principles of HyFlex which advocate the sharing of all materials with online students as being part of a *single learning community* (Beatty, 2019) and equivalent learning activities in all modes (Beatty, 2007).

### Cognitive engagement

Students in online classes may experience various engagement levels, ranging from boredom, interest, frustration, delight, confusion, and enthusiasm. Being able to identify these engagement levels is essential to providing students with personalized pedagogical support (Dewan et al., 2019). Students' behaviors are especially influenced by their cognitive engagement, which reflects their investment in learning to understand and master difficult concepts, participation in class, and effort to improve their academic performance (El-Sayad et al., 2021). Students' cognitive engagement determines their understanding and academic performance, but the willingness to listen, interact, concentrate and (actively) participate in the classroom should be intrinsic and must arise from students (Nagadeepa et al., 2021). Nevertheless, academics can help students by rethinking their assessment (for learning) practices, incorporating authentic course material into assignments, and interacting frequently with them (Walker & Koralesky, 2021). Academics teaching blended/hybrid courses should also take into consideration "the cognitive overload generated by instructors' split attention between online students and face-to-face students," which may have an impact on "instructors' attitudes towards online students" (Lakhal et al., 2020).

## Behavioral engagement

Students' participation, persistence, contribution, focus, efforts, and active involvement are essential to perform academic tasks (Nguyen et al., 2018). Several research studies, such as Liu et al. (2019), noted that students showing higher engagement rates achieved higher assessment results and demonstrated more autonomy in their learning. This engagement refers to emotional and cognitive aspects but also to behavior-related issues. Assessments often serve as motivators for students to engage with course materials and if there is a gap in assessment, students may feel demotivated or disengaged, as they may not see the value in actively participating in learning activities (Zhu et al., 2021). According to D'Mello (2021), behavioral engagement broadly refers to learners' participation in learning, including effort, persistence, and concentration. In other words, behavioral engagement pertains to students' investment in the learning task, such as how they allocate their effort towards learning and their understanding and mastery of the material (D'Mello, 2021). It also relates to how they interact with their classmates in, for instance, online group work (Knopf et al., 2021). Higher education institutions (HEIs) should therefore carefully monitor and analyze all indicators that relate to students' behavioral engagement, such as participation in campus life, achievement of intended learning outcomes, attendance, the effort to stay on task, contribution, participation in class/online discussions, involvement in academic activities, time spent on work, and perseverance when faced with challenging tasks (Bowden et al., 2021).

## Engagement in the HyFlex modality

Implementation of the HyFlex modality must provide an overall "flexible framework" (Wright, 2016, p. 88). Flexibility helps foster various levels of engagement for students through its synchronous or asynchronous learning options, and therefore, it must also be adaptive to the three dimensions of engagement (Fredricks et al., 2004). That is, it needs to foster positive connections between students' behavioral, emotional, and cognitive engagement within a course (Heilporn et al., 2021). Graduate students, due to the nature of their degree and discipline(s), may need to have access to content regardless of geographical location (when they do fieldwork, for instance) or time constraints (an internship overseas). As a result, students are more likely to engage with the material since they can fit learning into their schedules more easily. This accessibility enables a wider range of students to participate in learning activities, including those who may have other commitments such as jobs or family responsibilities. Additionally, this flexibility caters to individual learning (and navigation) styles and preferences, enabling students to spend more time on challenging concepts while moving quickly through material they already understand. This is often carried out through instructors identifying teaching expectations, demonstrating desired behavior for interactions in face-to-face and online environments, and ensuring student support through clear and timely feedback options (Lohmann et al., 2021). Binnewies and Wang (2019) have suggested these elements are commonly seen through teamwork and peer review

activities in HyFlex courses. Heilporn and Lakha (2021) further highlighted that such activities can engage face-to-face, synchronous, and asynchronous students at the same time while encouraging communication in the instructor/student relationship and providing a supportive learning structure.

Despite hesitations prior to COVID-19, universities worldwide have increasingly adopted hybrid modalities since blended teaching and learning positively affect students' performance. The literature points to several benefits of engagement with the HyFlex modality. Students enrolled in hybrid courses achieve similar or better outcomes than those enrolled in face-to-face courses (Green et al., 2018). Hybrid learning is especially useful when academics and students are engaged with innovation and when they focus on academic content rather than its technical aspects (Keržič et al., 2019). In the hybrid learning modality, academics and students establish better relations based on knowledge, performance, as well as emotions and social skills. In other words, hybrid modalities help reinforce students' skills in emotional intelligence (Iqbal et al., 2022). However, students value quality real-time interactions with academic staff (and personnel with teaching responsibilities), and asynchronous discussions and chats seem to negatively affect students' engagement and learning outcomes (Zhu et al., 2021). Koskinen (2018) found, for instance, that graduate students felt little connection with the "content, classmates, and the instructor" in courses that offered little to no interaction (Koskinen, 2018, p. 80).

The interaction equivalency theorem proposed by Anderson (2003) suggested that "deep and meaningful formal learning is supported as long as one of the three forms of interaction (student-teacher; student-student; student-content) is at a high level. The other two may be offered at minimal levels, or even eliminated, without degrading the educational experience" (p. 4). Universities should therefore encourage the integration of thoughtfully planned peer-hybrid learning activities (paragogy) into courses to help, for instance, students develop their critical thinking skills, as well as their engagement with the content (Khan & Iqbal, 2021). The use of various digital tools in hybrid courses such as apps, live chats, AI tools, automated (written/audio) ipsative authentic feedback, virtual teaching assistants or digital assessment may promote students' behavioral and emotional engagement. Academics can explore, experiment with and use different teaching and learning activities and strategies (Khan & Iqbal, 2021) such as interactive simulations, metaverse games, live polling (Houy, 2023), digital bulletin boards/online project management tools, AI-enhanced Video Based Learning (Shehata et al., 2023), online role-playing (RPG platforms, MMORPG *multiplayer online role-playing games*), flipped classroom, challenged-based projects, formative evaluations, interactive diagnostic quizzes or immersive learning tasks/virtual experiments, etc. All these may help provide a similar/comparable learning experience to all students to achieve equivalent learning outcomes. Among these initiatives, Coyne et al. (2018) argued that (short) simulated video resources were one of the most useful tools in a hybrid environment because they increased students' understanding of the academic content and enhanced their interactions with professors, at the

university and at home.

Hybrid initiatives determine students' motivation for learning and the more students are self-directed learners, the higher their acceptance of the technology is in blended or hybrid courses (Shimizu et al., 2019). Furthermore, to improve students' engagement with the online components of courses, universities should consider five main principles. *First, an integrated deliberate and reflective approach.* Academics need to consider several perspectives such as the design of the course (learning and teaching activities and tools) for the online modality, the learning outcomes, the accreditation standards, the exact role and responsibility of the teaching assistants, labs, cognitive load, autonomy and presence, multimodality, formative and summative assessment, team collaboration and monitoring, to propose unique and relevant content to their students (Hultberg et al., 2018; de Nooijer et al., 2021).

*Second, previous experience, data, and the untapped potential of social media.* Calonge et al. (2019) advocated that "purposeful learning and curriculum design decisions are a fundamental means to enhance participant engagement, motivation and performance in an online course" (p. 100). Academics need to harness data, analyze previous experience (their own and students') and satisfaction with courses facilitated online, reflect, generate actionable insights, and make decisions based on this information (Stephens et al., 2021).

*Third, equity, interactivity, and equivalency.* Some students may lack access to the necessary means, technology, tools, and bandwidth to fully partake in some aspects of the learning initiatives (Hines et al., 2020). A study by Platt et al. (2014) found that participants perceived fewer opportunities to interact with their instructor and classmates (p. 494) in the online mode. In terms of equivalency, participants did not see online courses as equivalent to face-to-face courses in a general sense.

*Fourth, students' mental health.* This issue has a great impact on well-being and academic performance, which determines students' enrolment and adaptation (and retention) to university life (Cobo-Rendón et al., 2020). A National Student Clearinghouse Research Center report indicated a 1.4 million university enrolment decline in the United States during the pandemic (NSCRC, 2022).

*Fifth, communication and mutual respect.* Effective, constant, transparent, and clear communication is essential, as is mutual respect among professors and students (Afzal, 2021) to create and sustain a safe, collegial environment.

### **Practical suggestions based on literature**

Based on the published literature, the RQ and the supplementary research question,

"Is a sense of belonging relevant to graduate students in HyFlex courses?"

The authors of this article propose several suggestions to improve emotional, cognitive, behavioral engagement and a sense of belonging in graduate HyFlex courses.

### **Practical suggestions to improve emotional engagement and sense of belonging at the graduate level**

To improve emotional engagement and sense of belonging, higher education institutions should consider the three following suggestions.

First, positive social interaction is important to student engagement and a sense of belonging. By deliberately designing and encouraging opportunities for social interaction, such as using synchronous class time to build peer networks, peer-to-peer, and student-to-instructor sharing activities (van Gijn-Grosvenor et al., 2020), emotional engagement can be fostered. Asynchronous activities to build positive social interaction can include the use of discussion boards, chatbots, and other interactive online tools such as quizzes and polling.

Second, to improve emotional engagement and sense of belonging when using Zoom or Teams and to avoid passivity, or a feeling that the student is "watching a lecture" rather than participating in an online class, it is important to use activities like chat functions and breakout rooms to allow students to not only discuss the content in small groups but to also build a sense of engagement and belonging and have their socio-emotional needs met (Saldanha et al., 2021).

Third, teaching staff should demonstrate their passion and emotional investment in their teaching, as well as encourage students to adopt "active to learn" behaviors (Mentzer et al., 2023) such as turning the camera on, unmuting in small group classes, and leading small breaks for movement (for example, a stretch break every 30 minutes, with brainstorming).

These strategies of modelling and promoting active and present participation can help facilitate positive engagement, which has been shown to increase emotional engagement and a sense of belonging (Peper et al., 2021). In fact, Deng (2021) indicated that emotional engagement was vastly more impactful on (online) student satisfaction than cognitive and behavioral engagement. This confirms previous work by Kucuk and Richardson (2019) who reported that data collected from 123 graduate students enrolled in an online program in the U.S. indicated that "emotional engagement was found to be one of the most important determining factors of satisfaction" (p. 207), which also highlights the importance of collecting empirical data.

### **Practical suggestions to help improve cognitive engagement at the graduate level**

Four initiatives may help reinforce students' cognitive engagement and mastery of core concepts at the graduate level.

First, HyFlex courses' learning and teaching activities and assessment tasks should be thoughtfully and purposely planned and designed to foster deep learning (Hultberg & Calonge, 2017), maximize interactivity and functioning knowledge, and provide an equivalent learning experience for students in synchronous (in person), synchronous (online) and asynchronous (online) modes (Casimiro, 2016; Hollister et al., 2022). As argued by Conrad et al. (2022), "shifting to online teaching is not simply duplicating the learning materials in the digital format" (p. 551).

Second, to increase motivation and interest, a variety of relevant tools and active learning strategies should be experimented with: MOOCs for skills (Calonge & Aman Shah, 2016), Open Educational Resources, live polling, videos and podcasts, games (e.g., <https://inworld.ai/studio>), chatbots (Calonge et al., 2023), AI tools (Kamalov et al., 2023a), collaborative lecture notes taking, interactive quizzes (e.g., Quizlet), online forums, open-access computer algebra systems (Kamalov et al., 2023b), community sharing initiatives and simulations (Ogunyemi et al., 2022).

Third, promote presence and planned online discussions to allow students to efficiently interact with professors and peers (Tang et al., 2021) and to evaluate students' cognitive engagement levels in these online forums (Kew & Tasir, 2021).

Fourth, encourage instructor-students' interactions and peer (online)-to-peer (classroom) teamwork and assessment (e.g., collaborative digital whiteboards) for students to ask questions, evaluate projects and work together (Wang, 2022), thus fostering an equivalent learning experience for all.

### **Practical suggestions to help improve behavioural engagement at the graduate level**

Higher education institutions should consider the following five suggestions to reinforce students' behavioral engagement.

First, universities should implement a holistic consideration of contexts, dimensions, factors and domains of engagement, and train academics and staff with teaching responsibilities to develop decision-making strategies for the early prevention of disengagement causes (Hasanov et al., 2021) using AI-powered predictive data analytics (Almusaed et al., 2023).

Second, universities should provide flexible learning options to graduate students in terms of time, place, and pace of learning (Kokoç, 2019).

Third, academics must deliberately design online courses that are relevant to graduate students in terms of level, content, and knowledge (Fabian et al., 2022), but also in terms of social support, online support, instructor presence, and management of students' anxiety (Bond & Bedenlier, 2019).

Fourth, academics should consider "pedagogy at the forefront of the design" (Zhang et al., 2022) and integrate cognitively engaging learning initiatives, which involve much more than simply presenting content through PowerPoint presentations and videos (Kennedy, 2020). A study by Houy (2023) indicated that polling and quiz slides had a motivating and engaging effect on students.

Fifth, academics need to design online activities that are compatible with multiple devices and formats and create tasks with real-life applications (Sugden et al., 2021), providing graduate-level students with greater opportunities to apply their course knowledge to their research and their professional lives.

### **Practical suggestions to implement HyFlex engagement strategies at the graduate level**

To attenuate the feelings of isolation described by Conrad et al. (2022) there is a need to provide equivalent learning experiences and to improve social presence in the same space.

Higher education institutions need to rethink and redesign infrastructures and learning spaces (interactive learning and collaboration spaces such as classrooms, meeting rooms and the library/learning commons) to support HyFlex (Detyna et al., 2023). These changes should be made to facilitate in-class and remote instructor-students interactions, and peer (online)-to-peer (classroom) real-time and breakout room interactions, teamwork, and assessment. A recent study by Calonge et al. (2023) details the implementation strategies used at Mohamed bin Zayed University of Artificial Intelligence (United Arab Emirates) to design HyFlex technology-enhanced classrooms. The authors provide examples of well-thought, modular, flexible, optimized classrooms using a human-centered (students and faculty) design-thinking/service design process with purposeful integration of technology to actively engage remote and F2F students synchronously and asynchronously. The article also provides qualitative feedback from both students and faculty highlighting the positive aspects of HyFlex learning space design in terms of flexibility, collaboration, and a sense of belonging.

### **Conclusion**

This study argues that as the HyFlex modality emphasizes blended learning principles and offers flexibility through its synchronous or asynchronous learning options, it is uniquely positioned for the current post-pandemic teaching and learning environment. Graduate students require emotional, cognitive, and behavioral engagement for optimal learning, and as argued, the Hyflex modality can provide it.

By combining face-to-face and online learning methods, HyFlex is adaptable to social distancing measures (if need be) and provides students flexibility and choice for how, where, and when (and with whom?) they engage with course content. To improve connection with the content, classmates, and the instructor, as well as encourage

meaningful interactions, cross-pollination of ideas, motivation in those mandatory graduate level courses, and increased emotional, cognitive, and behavioral engagement, the use of interdisciplinary, authentic and hybrid (research) challenge-based learning (CBL) team (F2F/remote) projects is recommended, using a Community of Inquiry (CoI) format (Garrison et al., 1999). Thus, the HyFlex teaching and learning modality may be able to support and encourage a sense of belonging, as well as emotional, cognitive, and behavioral engagement, while achieving equivalent learning outcomes for online and in-person students IF equivalent learning experiences are intentionally constructed (Simonson et al., 1999). When dealing with low-quality internet access, especially in contexts where students come from low socio-economic backgrounds in the Global North and South (including refugee contexts), integrating HyFlex for non-mandatory courses requires, however, careful consideration and adaptation to avoid exacerbating existing inequalities. No or limited access to real-time high-speed internet or Wi-Fi can pose significant accessibility challenges for students, hindering their ability to participate fully (or at all) in online activities. Ensuring recordings (and downloadable PDFs) are available for those who could not attend live sessions due to connectivity issues is key (Shah & Calonge, 2019; Shah & Calonge, 2023). Other downloadable lightweight resources should be ready so that students can access them offline, and local meetups or study groups for students facing connectivity challenges can be organized to collaborate and engage with course and research materials together.

To foster cognitive engagement, higher education can adopt principles from cognitive load theory applied across platforms, having instructional videos available at crucial points in the learning process and encouraging equivalent meaningful student interactions with both peers and instructors, in all modalities. Furthermore, as argued by Salas-Pilco et al. (2022), “educators from HEIs should be offered in-service training and professional development on the application of emerging technologies and the combination of technological and pedagogical skills to conduct learning activities that promote students’ behavioral, cognitive and affective engagement.” That is, there is a need for an increase in faculty development and support on HyFlex (Armstrong, 2022), Learning Experience (LX) design (Howell et al., 2023), and Universal Design for Learning (UDL).

## References

Abdelmalak, M. M. M., & Parra, J. L. (2016). Expanding learning opportunities for graduate students with HyFlex course design. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, 6(4), 19-37. <https://doi.org/10.4018/IJOPCD.2016100102>

Afzal, M. H. B. (2021). Recommended practices for doctoral students in navigating and engaging in online courses during COVID-19 pandemic: A personal narrative from a doctoral candidate. *Biochemistry and Molecular Biology Education*, 49(5), 679-680. <https://doi.org/10.1002/bmb.21550>

Almusaed, A., Almssad, A., Yitmen, I., & Homod, R. Z. (2023). Enhancing student engagement: Harnessing “AIED”’s power

in hybrid education—A review analysis. *Education Sciences*, 13(7), 632. <https://doi.org/10.3390/educsci13070632>

Anderson, T. (2003). Getting the mix right again: An updated and theoretical rationale for interaction. *The International Review of Research in Open and Distributed Learning*, 4(2). <https://doi.org/10.19173/irrodl.v4i2.149>

Armstrong, E. D. (2022). *Gaps in professional development and knowledge of teaching HyFlex courses in higher education* [Doctoral dissertation, University of the Southwest]. <https://eric.ed.gov/?id=ED621726>

Beatty, B. J. (2007). Transitioning to an online world: Using HyFlex courses to bridge the gap. In C. Montgomerie & J. Seale (Eds.), *Proceedings of ED-MEDIA 2007--World conference on educational multimedia, hypermedia & telecommunications* (pp. 2701–2706). Vancouver: Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/25752/>

Beatty, B. J. (2019). Teaching a hybrid-flexible course. In B. J. Beatty (Ed.), *Hybrid-Flexible course design*. <https://dx.doi.org/10.59668/33>

Binnewies, S., & Wang, Z. (2019). Challenges of student equity and engagement in a HyFlex Course. In *Blended learning designs in STEM higher education* (pp. 209-230). Springer, Singapore. [https://doi.org/10.1007/978-981-13-6982-7\\_12](https://doi.org/10.1007/978-981-13-6982-7_12)

Bockorny, K. M., Giannavola, T. M., Mathew, S., & Walters, H. D. (2023). Effective engagement strategies in HyFlex modality based on intrinsic motivation in students. *Active Learning in Higher Education*. <https://doi.org/10.1177/14697874231161364>

Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational technology: Towards a conceptual framework. *Journal of Interactive Media in Education*, 2019(1). <https://doi.org/10.5334/jime.528>

Bowden, J. L. H., Tickle, L., & Naumann, K. (2021). The four pillars of tertiary student engagement and success: A holistic measurement approach. *Studies in Higher Education*, 46(6), 1207-1224. <https://doi.org/10.1080/03075079.2019.1672647>

Bozan, K., Gaskin, J., & Stoner, C. (2023). Student engagement in the HyFlex and online classrooms: Lessons from the COVID-19 pandemic. *Technology, Knowledge and Learning*, 1-28. <https://doi.org/10.1007/s10758-023-09661-x>

Burke, A. (2019). Student retention models in higher education: A literature review. *College and University*, 94(2), 12-21. <https://www.proquest.com/scholarly-journals/student-retention-models-higher-education/docview/2232610556/se-2>

Calonge, D. S., & Aman Shah, M. (2016). MOOCs, graduate skills gaps, and employability: A qualitative systematic review of the literature. *International Review of Research in Open and Distributed Learning: IRRDL*, 17(5), 67-90. <https://doi.org/10.19173/irrodl.v17i5.2675>

- Calonge, D. S., Connor, M., Hultberg, P., Shah, M. A., & Aguerrebere, P. M. (2022). Contactless higher education: A SWOT analysis of emergency remote teaching and learning during COVID-19. *Journal of Educational Studies and Multidisciplinary Approaches*, 2(1). <https://doi.org/10.51383/jesma.2022.22>
- Calonge, D. S., Riggs, K. M., Shah, M. A., & Cavanagh, T. A. (2019). Using learning analytics to improve engagement, learning, and design of Massive Open Online Courses. In *Fostering multiple levels of engagement in higher education environments* (pp. 76-107). IGI Global. <http://dx.doi.org/10.4018/978-1-5225-7470-5.ch004>
- Calonge, D. S., Shah, M. A., Aguerrebere, P. M., Abdulla, N., Connor, M., Badr, M., & Blakemore, E. (2023). Should I stay or should I go? International students' challenges and opportunities to secure employment in their host country after graduation. A scoping review using PRISMA. *Journal of Applied Learning and Teaching*, 6(2), 187-201. <https://doi.org/10.37074/jalt.2023.6.2.20>
- Calonge, D. S., Smail, L., & Kamalov, F. (2023). Enough of the chit-chat: A comparative analysis of four AI chatbots for calculus and statistics. *Journal of Applied Learning and Teaching*, 6(2), 346-357. <https://doi.org/10.37074/jalt.2023.6.2.22>
- Calonge, D. S., Thompson, M., Hassock, L., & Yaqub, M. (2023). Hybrid Flexible (HyFlex) learning space design and implementation at graduate level: An iterative process. *Cogent Education*, 10(2), 2277001. <https://doi.org/10.1080/2331186X.2023.2277001>
- Casimiro, L. T. (2016). Cognitive engagement in online intercultural interactions: Beyond analytics. *International Journal of Information and Education Technology*, 6(6), 441-447. <https://doi.org/10.7763/IJiet.2016.V6.729>
- Chen, H., & Yang, M. (2022). Online student response systems and student engagement in large EFL classrooms. *Journal of Applied Learning and Teaching*, 5(1), 60-70. <https://doi.org/10.37074/jalt.2022.5.1.3>
- Cobo-Rendón, R., Pérez-Villalobos, M. V., Páez-Rovira, D., & Gracia-Leiva, M. (2020). A longitudinal study: Affective wellbeing, psychological wellbeing, self-efficacy and academic performance among first-year undergraduate students. *Scandinavian Journal of Psychology*, 61(4), 518-526. <https://doi.org/10.1111/sjop.12618>
- Conrad, C., Deng, Q., Caron, I., Shkurska, O., Skerrett, P., & Sundararajan, B. (2022). How student perceptions about online learning difficulty influenced their satisfaction during Canada's COVID-19 response. *British Journal of Educational Technology*, 53(3), 534-557. <https://doi.org/10.1111/bjet.13206>
- Coyne, E., Frommolt, V., Rands, H., Kain, V., & Mitchell, M. (2018). Simulation videos presented in a blended learning platform to improve Australian nursing students' knowledge of family assessment. *Nurse Education Today*, 66, 96-102. <https://doi.org/10.1016/j.nedt.2018.04.012>
- Cunningham, U. (2014). Teaching the disembodied: Othering and activity systems in a blended synchronous learning situation. *International Review of Research in Open and Distributed Learning*, 15(6), 33-51. <https://doi.org/10.19173/irrodl.v15i6.1793>
- de Nooijer, J., Schneider, F., & Verstegen, D. M. (2021). Optimizing collaborative learning in online courses. *The Clinical Teacher*, 18(1), 19-23. <https://doi.org/10.1111/tct.13243>
- Deng, R. (2021). Emotionally engaged learners are more satisfied with online courses. *Sustainability*, 13(20), 11169. <https://doi.org/10.3390/su132011169>
- Detyna, M., Sanchez-Pizani, R., Giampietro, V., Dommett, E. J., & Dyer, K. (2023). Hybrid flexible (HyFlex) teaching and learning: Climbing the mountain of implementation challenges for synchronous online and face-to-face seminars during a pandemic. *Learning Environments Research*, 26(1), 145-159. <https://doi.org/10.1007/s10984-022-09408-y>
- Dewan, M., Murshed, M., & Lin, F. (2019). Engagement detection in online learning: A review. *Smart Learning Environments*, 6(1), 1-20. <https://doi.org/10.1186/s40561-018-0080-z>
- D'Mello, S. K. (2021). Improving student engagement in and with digital learning technologies. In *OECD digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots* (p. 79). <https://doi.org/10.1787/8a451974-en>
- Durón-Ramos, M. F., & García-Vázquez, F. (2018). Orientation to happiness as predictor of university students' engagement. *International Journal of Evaluation and Research in Education*, 7(4), 294-298. <https://doi.org/10.11591/ijere.v7i4.15446>
- El-Sayad, G., Md Saad, N. H., & Thurasamy, R. (2021). How higher education students in Egypt perceived online learning engagement and satisfaction during the COVID-19 pandemic. *Journal of Computers in Education*, 8(4), 527-550. <https://doi.org/10.1007/s40692-021-00191-y>
- Fabian, K., Smith, S., Taylor-Smith, E., & Meharg, D. (2022). Identifying factors influencing study skills engagement and participation for online learners in higher education during COVID-19. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13221>
- Fernández-García, C. M., Rodríguez-Álvarez, M., & Viñuela-Hernández, M. P. (2021). University students and their perception of teaching effectiveness. Effects on students' engagement. *Revista de Psicodidáctica (English ed.)*, 26(1), 62-69. <https://doi.org/10.1016/j.psicoe.2020.11.006>
- Fook, J. (2011). Developing critical reflection as a research method. In J. Higgs, A. Titchen, D. Horsfall, & D. Bridges (Eds.), *Creative spaces for qualitative researching: Living research* (5th ed., pp. 55-64). Sense Publishers. [https://doi.org/10.1007/978-94-6091-761-5\\_6](https://doi.org/10.1007/978-94-6091-761-5_6)
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School

- engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- García-Martínez, I., Landa, J. M. A., & León, S. P. (2021). The mediating role of engagement on the achievement and quality of life of university students. *International Journal of Environmental Research and Public Health*, 18(12), 6586. <https://doi.org/10.3390/ijerph18126586>
- Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, 2(2-3), 87-105. [https://doi.org/10.1016/S1096-7516\(00\)00016-6](https://doi.org/10.1016/S1096-7516(00)00016-6)
- Gibbs, G. (1988). *Learning by doing: A guide to teaching and learning methods*. Further Education Unit. <https://thoughtsmostlyaboutlearning.files.wordpress.com/2015/12/learning-by-doing-graham-gibbs.pdf>
- Green, R. A., Whitburn, L. Y., Zacharias, A., Byrne, G., & Hughes, D. L. (2018). The relationship between student engagement with online content and achievement in a blended learning anatomy course. *Anatomical Sciences Education*, 11(5), 471-477. <https://doi.org/10.1002/ase.1761>
- Hasanov, Z., Antoniou, P., Suleymanov, E., & Garayev, V. (2021). The impact of behavioral, cognitive, and emotional dimensions of student engagement on student learning: The case of Azerbaijani higher education institutions. *International Journal of Knowledge and Learning*, 14(1), 10-38. <http://dx.doi.org/10.1504/IJKL.2021.10037816>
- Heilporn, G., & Lakhal, S. (2021). Converting a graduate-level course into a HyFlex modality: What are affective engagement strategies? *The International Journal of Management Education*, 19(1). <https://doi.org/10.1016/j.ijme.2021.100454>
- Heilporn, G., Lakhal, S., & Bélisle, M. (2021). An examination of teachers' strategies to foster student engagement in blended learning in higher education. *International Journal of Educational Technology in Higher Education*, 18(25), 1-25. <https://doi.org/10.1186/s41239-021-00260-3>
- Hickson, H. (2015). Becoming a critical narrativist: Using critical reflection and narrative inquiry as research methodology. *Qualitative Social Work*, 15(3), 380-391. <https://doi.org/10.1177/1473325015617344>
- Hines, S. L., Vedral, A. J., Jefferson, A. E., Drymon, J. M., Woodrey, M. S., Mabey, S. E., & Sparks, E. L. (2020). Engaging online students by activating ecological knowledge. *Ecology and Evolution*, 10(22), 12472-12481. <https://doi.org/10.1002/ece3.6739>
- Hollister, B., Nair, P., Hill-Lindsay, S., & Chukoskie, L. (2022, May). Engagement in online learning: Student attitudes and behavior during COVID-19. *Frontiers in Education*, 7, 851019. Frontiers Media SA. <https://doi.org/10.3389/educ.2022.851019>
- Holzweiss, P. C., Joyner, S. A., Fuller, M. B., Henderson, S., & Young, R. (2014). Online graduate students' perceptions of best learning experiences. *Distance Education*, 35(3), 311-323. <https://doi.org/10.1080/01587919.2015.955262>
- Houy, Y. (2023). *Engaging HyFlex: Bridging the remote/in-person engagement gap through instant polling slides*. [https://digitalscholarship.unlv.edu/btp\\_expo/194/](https://digitalscholarship.unlv.edu/btp_expo/194/)
- Howell, E., Hubbard, K., Linder, S., Madison, S., Ryan, J., & Bridges, W. C. (2023). HyFlex pedagogy: Six strategies supported by design-based research. *Journal of Applied Research in Higher Education*. <https://doi.org/10.1108/JARHE-02-2023-0050>
- Hultberg, P., & Calonge, D. S. (2017). Effective teaching of economics: A constrained optimization problem. *Journal of Economic Education*, 48(4), 265-275. <https://doi.org/10.1080/00220485.2017.1353458>
- Hultberg, P., Calonge, D. S., & Lee, A. E. S. (2018). Promoting long-lasting learning through instructional design. *Journal of the Scholarship of Teaching and Learning*, 18(3). <https://doi.org/10.14434/josotl.v18i3.23179>
- Iqbal, J., Asghar, M. Z., Ashraf, M. A., & Yi, X. (2022). The impacts of emotional intelligence on students' study habits in blended learning environments: The mediating role of cognitive engagement during COVID-19. *Behavioral Sciences*, 12(1), 14. <https://doi.org/10.3390/bs12010014>
- Kamalov, F., Calonge, D. S., & Gurrib, I. (2023a). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability*, 15(16), 12451. <https://doi.org/10.3390/su151612451>
- Kamalov, F., Calonge, D. S., Leung, H. H., Johnson, J., & El Khatib, Z. (2023b, May). Leveraging computer algebra systems in calculus: A case study with SymPy. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-6). IEEE. <https://doi.org/10.1109/EDUCON54358.2023.10125196>
- Kennedy, G. (2020). *What is student engagement in online learning... and how do I know when it is there?* [Discussion Papers, Melbourne Centre for the Study of Higher Education]. University of Melbourne. [https://melbourne-cshe.unimelb.edu.au/\\_\\_data/assets/pdf\\_file/0004/3362125/student-engagement-online-learning\\_final.pdf](https://melbourne-cshe.unimelb.edu.au/__data/assets/pdf_file/0004/3362125/student-engagement-online-learning_final.pdf)
- Keržič, D., Tomažević, N., Aristovnik, A., & Umek, L. (2019). Exploring critical factors of the perceived usefulness of blended learning for higher education students. *PLoS One*, 14(11), e0223767. <https://doi.org/10.1371/journal.pone.0223767>
- Kew, S. N., & Tasir, Z. (2021). Analysing students' cognitive engagement in e-learning discussion forums through content analysis. *Knowledge Management & E-Learning: An International Journal*, 13(1), 39-57. <https://doi.org/10.34105/j.kmel.2021.13.003>
- Khan, S., & Iqbal, S. (2021). Innovative ways of student engagement for active learning in science courses of nursing in the four-year baccalaureate program. *Journal*

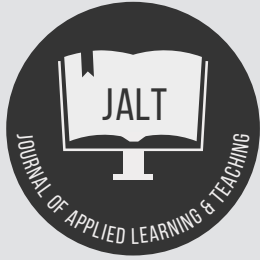
- of the Pakistan Medical Association, 1-10. <https://doi.org/10.47391/JPMA.308>
- Knopf, T., Stumpp, S., & Michelis, D. (2021, July). How online collaborative learning leads to improved online learning experience in higher education. In *Proceedings of the 8th European conference on social media*. [https://www.researchgate.net/publication/352888884\\_How\\_Online\\_Collaborative\\_Learning\\_leads\\_to\\_improved\\_Online\\_Learning\\_Experience\\_in\\_Higher\\_Education](https://www.researchgate.net/publication/352888884_How_Online_Collaborative_Learning_leads_to_improved_Online_Learning_Experience_in_Higher_Education)
- Kohnke, L., & Moorhouse, B. L. (2021). Adopting HyFlex in higher education in response to COVID-19: Students' perspectives. *Open Learning: The Journal of Open, Distance And E-Learning*, 36(3), 231-244. <https://doi.org/10.1080/02680513.2021.1906641>
- Kokoç, M. (2019). Flexibility in e-learning: Modelling its relation to behavioural engagement and academic performance. *Themes in eLearning*, 12(12), 1-16. <https://files.eric.ed.gov/fulltext/EJ1251161.pdf>
- Koskinen, M. (2018). *Understanding the needs of adult graduate students: An exploratory case study of a HyFlex learning environment* [Doctoral dissertation, Northeastern University]. NU repository. <https://repository.library.northeastern.edu/files/neu:m044c738v/fulltext.pdf>
- Kuchinski-Donnelly, D., & Krouse, A. M. (2020). Predictors of emotional engagement in online graduate nursing students. *Nurse Educator*, 45(4), 214-219. <https://doi.org/10.1097/NNE.0000000000000769>
- Kucuk, S., & Richardson, J. C. (2019). A structural equation model of predictors of online learners' engagement and satisfaction. *Online Learning*, 23(2), 196-216. <https://doi.org/10.24059/olj.v23i2.1455>
- Lakhal, S., Mukamurera, J., Bédard, M. E., Heilporn, G., & Chauret, M. (2020). Features fostering academic and social integration in blended synchronous courses in graduate programs. *International Journal of Educational Technology in Higher Education*, 17(1), 1-22. <https://doi.org/10.1186/s41239-020-0180-z>
- Lapsley, R., Kulik, B., Moody, R., & Arbaugh, J. B. (2008). Is identical really identical? An investigation of equivalency theory and online learning. *Journal of Educators Online*, 5(1), n1. <http://dx.doi.org/10.9743/JEO.2008.1.3>
- Lim, D. H., Morris, M. L., & Kupritz, V. W. (2007). Online vs. blended learning: Differences in instructional outcomes and learner satisfaction. *Journal of Asynchronous Learning Networks*, 11(2), 27-42. <https://files.eric.ed.gov/fulltext/EJ842695.pdf>
- Liu, C., Lim, R., Taylor, S., & Calvo, R. A. (2019). Students' behavioural engagement in reviewing their tele-consultation feedback within an online clinical communication skills platform. *Computers in Human Behavior*, 94, 35-44. <https://doi.org/10.1016/j.chb.2019.01.002>
- Lobos Peña, K., Bustos-Navarrete, C., Cobo-Rendón, R., Fernández Branada, C., Bruna Jofré, C., & Maldonado Trapp, A. (2021). Professors' expectations about online education and its relationship with characteristics of university entrance and students' academic performance during the COVID-19 pandemic. *Frontiers in Psychology*, 12, 882. <https://doi.org/10.3389/fpsyg.2021.642391>
- Lohmann, M. J., Randolph, K. M., & Oh, J. H. (2021). Classroom management strategies for Hyflex instruction: Setting students up for success in the hybrid environment. *Early Childhood Education Journal*, 49, 807-814. <https://doi.org/10.1007/s10643-021-01201-5>
- Martín-Hernández, P., Gil-Lacruz, M., Gil-Lacruz, A. I., Azkue-Beteta, J. L., Lira, E. M., & Cantarero, L. (2021). Fostering university students' engagement in teamwork and innovation behaviors through Game-Based Learning (GBL). *Sustainability*, 13(24), 13573. <https://doi.org/10.3390/su132413573>
- McDavid, J. C., & Shepherd, R. P. (2021). Offering graduate evaluation degrees online: Comparing student engagement in two Canadian programs. *Canadian Journal of Program Evaluation*, 35(3), 363-373. <https://doi.org/10.3138/cjpe.69751>
- Mentzer, N., Krishna, B., Kotangale, A., & Mohandas, L. (2023). HyFlex environment: Addressing students' basic psychological needs. *Learning Environments Research*, 26(1), 271-289. <https://doi.org/10.1007/s10984-022-09431-z>
- Nagadeepa, C., Reenu, M., Raji, S., & Raja Kamal, C. H. (2021). Students' understanding and learning: Mediation effects of cognitive engagement in online classes. *Turkish Journal of Computer and Mathematics Education*, 12, 2932-2939. <https://doi.org/10.17762/turcomat.v12i6.6077>
- Nelson, T. A., Berg, E. A., Wood, N., & Hill, B. (2022). Student engagement in HyFlex courses during the COVID-19 pandemic. *Journal of College Student Development*, 63(1), 101-105. <https://doi.org/10.1353/csd.2022.0001>
- Nguyen, T. D., Cannata, M., & Miller, J. (2018). Understanding student behavioral engagement: Importance of student interaction with peers and teachers. *The Journal of Educational Research*, 111(2), 163-174. <https://doi.org/10.1080/00220671.2016.1220359>
- NSCRC (National Student Clearinghouse Research Center). (2022). *Overview: Spring 2022 enrollment estimates*. [https://nscresearchcenter.org/wp-content/uploads/CTEE\\_Report\\_Spring\\_2022.pdf](https://nscresearchcenter.org/wp-content/uploads/CTEE_Report_Spring_2022.pdf)
- Ogunyemi, A. A., Quaiocoe, J. S., & Bauters, M. (2022). Indicators for enhancing learners' engagement in massive open online courses: A systematic review. *Computers and Education Open*, 100088. <https://doi.org/10.1016/j.caeo.2022.100088>
- Oviawe, J. I. (2020). Technical education lecturers' knowledge of students' engagement in application of interactive instructional strategies. *Journal of Technology and Humanities*, 1(1), 1-10. <https://doi.org/10.53797/jthkkss>



- Peacock, S., Cowan, J., Irvine, L., & Williams, J. (2020). An exploration into the importance of a sense of belonging for online learners. *International Review of Research in Open and Distributed Learning*, 21(2), 18-35. <https://doi.org/10.19173/irrodl.v20i5.4539>
- Pechenkina, E., Laurence, D., Oates, G., Eldridge, D., & Hunter, D. (2017). Using a gamified mobile app to increase student engagement, retention and academic achievement. *International Journal of Educational Technology in Higher Education*, 14(1), 1-12. <https://doi.org/10.1186/s41239-017-0069-7>
- Pedler, M. L., Willis, R., & Nieuwoudt, J. E. (2021). A sense of belonging at university: Student retention, motivation, and enjoyment. *Journal of Further and Higher Education*, 1-12. <https://doi.org/10.1080/0309877X.2021.1955844>
- Peper, E., Wilson, V., Martin, M., Rosegard, & Harvey, R. (2021). Avoid Zoom fatigue, be present and learn. *NeuroRegulation*, 8(1), 47-56. <https://doi.org/10.15540/nr.8.1.47>
- Platt, C. A., Amber, N. W., & Yu, N. (2014). Virtually the same? Student perceptions of the equivalence of online classes to face-to-face classes. *Journal of Online Learning and Teaching*, 10(3), 489. [https://www.researchgate.net/publication/271205198\\_Virtually\\_the\\_same\\_Student\\_perceptions\\_of\\_the\\_equivalence\\_of\\_online\\_classes\\_vs\\_face-to-face\\_classes](https://www.researchgate.net/publication/271205198_Virtually_the_same_Student_perceptions_of_the_equivalence_of_online_classes_vs_face-to-face_classes)
- Raes, A., Detienne, L., Windey, I., & Depaeppe, F. (2020). A systematic literature review on synchronous hybrid learning: Gaps identified. *Learning Environments Research*, 23(3), 269-290. <https://doi.org/10.1007/s10984-019-09303-z>
- Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performance in an online module: Implications for institutional e-learning policy. *Education and Information Technologies*, 26(3), 2623-2656. <https://doi.org/10.1007/s10639-020-10375-1>
- Romero-Hall, E., & Ripine, C. (2021). Hybrid flexible instruction: Exploring faculty preparedness. *Online Learning Journal*, 25(3), 289-312. <https://doi.org/10.24059/olj.v25i3.2426>
- Salas-Pilco, S. Z., Yang, Y., & Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: A systematic review. *British Journal of Educational Technology*, 53(3), 593-619. <https://doi.org/10.1111/bjjet.13190>
- Saldanha, K., Currin-McCulloch, J., Muskat, B., Simon, S. R., Bergart, A. M., Mesbur, E. S., Guy, D., Chilwalo, N. B., Seck, M. M., Tully, G., Lind, K., Lee, C. D., Hall, N., & Kelly, D. (2021). Turning boxes into supportive circles: Enhancing online group work teaching during the COVID-19 pandemic. *Social Work with Groups*, 44(4), 310-327. <https://doi.org/10.1080/1609513.2021.1910110>
- Shah, M. A., & Calonge, D. S. (2019). Frugal MOOCs: An adaptable contextualized approach to MOOC designs for refugees. *International Review of Research in Open and Distributed Learning*, 20(5), 1-19. <https://doi.org/10.19173/irrodl.v20i4.3350>
- Shah, M. A., & Calonge, D. S. (2023). Refugees' experiences with online higher education: Impact and implications through the pandemic. *Journal of Applied Learning and Teaching*, 6(1), 209-221. <https://doi.org/10.37074/jalt.2023.6.1.21>
- Shehata, S., Calonge, D. S., Purnell, P., & Thompson, M. (2023, July). Enhancing video-based learning using knowledge tracing: Personalizing students' learning experience with ORBITS. In *Proceedings of the 18th workshop on innovative use of NLP for Building Educational Applications (BEA 2023)* (pp. 100-107). <https://doi.org/10.18653/v1/2023.bea-1.8>
- Shimizu, I., Nakazawa, H., Sato, Y., Wolfhagen, I. H., & Könings, K. D. (2019). Does blended problem-based learning make Asian medical students active learners?: A prospective comparative study. *BMC Medical Education*, 19(1), 1-9. <https://doi.org/10.1186/s12909-019-1575-1>
- Simonson, M., Schlosser, C., & Hanson, D. (1999). Theory and distance education: A new discussion. *American Journal of Distance Education*, 13(1), 60-75. <https://doi.org/10.1080/08923649909527014>
- Stephens, J. B., & Morse, R. S. (2022). *Enhancing sense of belonging and satisfaction among online students in multi-track public affairs programs: A case analysis of immersion courses*. Teaching Public Administration. <https://doi.org/10.1177/01447394221076344>
- Stephens, J. M., Watson, P. W. S. J., Alansari, M., Lee, G., & Turnbull, S. M. (2021). Can online academic integrity instruction affect university students' perceptions of and engagement in academic dishonesty? Results from a natural experiment in New Zealand. *Frontiers in Psychology*, 12, 366. <https://doi.org/10.3389/fpsyg.2021.569133>
- Strayhorn, T. L. (2018). *College students' sense of belonging: A key to educational success for all students*. Routledge. <https://doi.org/10.4324/9781315297293>
- Subramainan, L., & Mahmoud, M. A. (2020). A systematic review on students' engagement in classroom: Indicators, challenges and computational techniques. *International Journal of Advanced Computer Science and Applications*, 11(1). <https://doi.org/10.14569/IJACSA.2020.0110113>
- Sugden, N., Brunton, R., MacDonald, J., Yeo, M., & Hicks, B. (2021). Evaluating student engagement and deep learning in interactive online psychology learning activities. *Australasian Journal of Educational Technology*, 37(2), 45-65. <https://doi.org/10.14742/ajet.6632>
- Swanson, E., Culver, K. C., Cole, D., & Rivera, G. (2021). Promoting at-risk student success in 4-year universities: Recommendations from the Thompson scholars learning communities. *Journal of Diversity in Higher Education*, 14(4), 457-462. <https://doi.org/10.1037/dhe0000345>

- Tang, Y. M., Chen, P. C., Law, K. M., Wu, C. H., Lau, Y. Y., Guan, J., ... & Ho, G. T. (2021). Comparative analysis of student's live online learning readiness during the coronavirus (COVID-19) pandemic in the higher education sector. *Computers & Education, 168*, 104211. <https://doi.org/10.1016/j.compedu.2021.104211>
- Thomas, L., Herbert, J., & Teras, M. (2014). A sense of belonging to enhance participation, success and retention in online programs. *The International Journal of the First Year in Higher Education, 5*(2), 69-80. <https://doi.org/10.5204/intjfyhe.v5i2.233>
- Thompson, S., & Thompson, N. (2018). *The critically reflective practitioner*. Bloomsbury Publishing.
- Trotter, J., & Qureshi, F. (2023). Student perspectives of hybrid delivery in a transnational education context during Covid-19. *Journal of Applied Learning & Teaching, 6*(1), 103-111. <https://doi.org/10.37074/jalt.2023.6.1.10>
- van Gijn-Grosvenor, E. L., & Huisman, P. (2020). A sense of belonging among Australian university students. *Higher Education Research & Development, 39*(2), 376-389. <https://doi.org/10.1080/07294360.2019.1666256>
- Walker, K. A., & Koralesky, K. E. (2021). Student and instructor perceptions of engagement after the rapid online transition of teaching due to COVID-19. *Natural Sciences Education, 50*(1). <https://doi.org/10.1002/nse2.20038>
- Wang, Q., & Huang, C. (2018). Pedagogical, social and technical designs of a blended synchronous learning environment. *British Journal of Educational Technology, 49*(3), 451-462. <https://doi.org/10.1111/bjet.12558>
- Wang, Y. (2022). Effects of teaching presence on learning engagement in online courses. *Distance Education, 43*(1), 139-156. <https://doi.org/10.1080/01587919.2022.2029350>
- Wilson, T. J., & Alexander, M. (2021). HyFlex course delivery: Addressing the change in course modality brought on by the pandemic. *Journal of the International Society for Teacher Education, 25*(2), 41-58. <https://doi.org/10.26522/jiste.v25i2.3668>
- Wright, D. (2016). The HyFlex course design: A case study on adult and career education courses. *National Social Science Journal, 48*(2), 88-93. <http://dx.doi.org/10.1007/s10984-022-09408-y>
- Xiao, M. (2021). Chinese international graduate students at Canadian universities: Language barriers, cultural identities, and perceived problems of engagement. *International Journal of Inclusive Education, 1-18*. <https://doi.org/10.1080/13603116.2021.1941318>
- Yang, D., Lavonen, J. M., & Niemi, H. (2018). Online learning engagement: Factors and results-evidence from literature. *Themes in eLearning, 11*(1), 1-22. <https://files.eric.ed.gov/fulltext/EJ1204753.pdf>
- Zepke, N., & Leach, L. (2010). Improving student engagement: Ten proposals for action. *Active Learning in Higher Education, 11*(3), 167-177. <https://doi.org/10.1177/1469787410379680>
- Zhang, L., Carter Jr, R. A., Qian, X., Yang, S., Rujimora, J., & Wen, S. (2022). Academia's responses to crisis: A bibliometric analysis of literature on online learning in higher education during COVID-19. *British Journal of Educational Technology, 53*(3), 620-646. <https://doi.org/10.1111/bjet.13191>
- Zhu, M., Berri, S., & Zhang, K. (2021). Effective instructional strategies and technology use in blended learning: A case study. *Education and Information Technologies, 26*(5), 6143-6161. <https://doi.org/10.1007/s10639-021-10544-w>
- Zou, C., Li, P., & Jin, L. (2021). Online college English education in Wuhan against the COVID-19 pandemic: Student and teacher readiness, challenges, and implications. *PloS One, 16*(10), e0258137. <https://doi.org/10.1371/journal.pone.0258137>

Copyright: © 2024. David Santandreu Calonge, Mariam Aman Shah, Melissa Connor, Patrik Hultberg, and Pablo Medina Aguerrebere. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Student engagement and fostering ownership of learning

---

Ajay Khatter <sup>A</sup>	A	Senior Lecturer, William Angliss Institute, Melbourne, Australia
Krishna Thalaachawr <sup>B</sup>	B	Lecturer, William Angliss Institute, Melbourne, Australia
Mark Blyth <sup>C</sup>	C	Lecturer, William Angliss Institute, Melbourne, Australia

---

### Keywords

---

Action research;  
assessments;  
class activities;  
student engagement;  
student motivation.

---

### Correspondence

---

Ajay.khatter@angliss.edu.au <sup>A</sup>

---

### Article Info

Received 18 December 2023  
Received in revised form 29 March 2024  
Accepted 17 April 2024  
Available online 19 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.38>

### Abstract

---

This research was undertaken to find pedagogical solutions that help teachers design engaging learning activities and assessments. These solutions aim to increase student engagement, encourage them to take responsibility for their learning in practical training and education settings and effect positive classroom change. An action research methodology was utilised for the researchers to learn through action, increase student engagement, and develop professionally and personally. Action research is a family of research methodologies that pursue action (or change) and research (or understanding) simultaneously using a cyclic or spiral process that alternates between action and critical reflection. This methodology involved action, evaluation, and reflection to gather evidence to implement classroom change. The research found that adopting student-centred teaching practices, such as project-based learning and collaborative activities, has substantially improved student involvement with the learning process. The students who participated in the activities showed increased interest, motivation, and active participation in their educational experiences.

## Introduction

This research was conducted to generate knowledge on increasing student engagement, ensuring they take ownership of their learning, and improving the educational process. The action research methodology systematically approached changing teaching and learning environments to affect classroom change positively. Action research generates knowledge about educational theory and practices and has been a powerful tool for improving learning and teaching. It has assisted in understanding which teaching and pedagogical strategies best increase student engagement and encourage them to take ownership of their learning (Alpert et al., 2022).

Education professionals review their instructional strategies and explore tactics that increase student engagement. A positive correlation exists between increased student engagement and student achievement (Pudjiarti et al., 2023). It is acknowledged that student engagement is a significant factor in academic performance (Wheaton, 2021). The in-class lecture continues as the predominant instructional strategy in most classrooms. Most teachers still rely on this time-honoured method of delivering information to their students. However, many researchers believe that classroom lectures are an inefficient method of helping students acquire the necessary knowledge and skills, and they criticise it for this reason (Gilboy et al., 2015). Research into disengaged students identified that this transmission model of instruction induced passivity and boredom in the students (Liu et al., 2023; Marks, 2000). In contrast, student-centred learning is where teachers function more as coaches than lecturers, allowing students a say in their education and personalising instruction so they can learn at their own pace (Kamali, 2023). As a student's educational journey advances, the teacher's role should transform from being a mere provider of knowledge to assuming the responsibilities of a facilitator and mentor.

When lectures are the only mode of instruction for a class, this form of passive learning steals valuable classroom time that could be better spent stimulating students' thinking, directing them towards finding solutions to real-world problems, and encouraging the direct application of concepts through active learning while the teacher is present (Abdullah et al., 2019). Individuals are responsible for their knowledge construction and reconstruction, which involves making sense of the latest information in light of what they may already be familiar with. Utilising active learning strategies is the most effective method for facilitating students' participation in knowledge construction and reconstruction. In academic literature, student engagement has been referred to as 'student experience,' 'academic engagement,' 'academic integration,' or 'student involvement' (Bowden et al., 2021). Student engagement can be defined as "A student's positive social, cognitive, emotional, and behavioural investments made when interacting with their tertiary institution and its focal agents (such as peers, employees and the institution itself)" (Bowden et al., 2021, p. 5).

Deep learning proponents believe that teachers are needed to boost student engagement. They state:

"High-quality teaching implies recognising that students must be engaged with the content of learning tasks in a way that is likely to enable them to reach understanding ... Sharp engagement, imaginative inquiry and the finding of a suitable level and style are all more likely to occur if teaching methods that necessitate student energy, problem-solving and cooperative learning are employed" (Ramsden, 2003, p. 97).

This study sought to explore Ramsden's claims by using action research to investigate and trial improvements in pedagogical strategies to enhance student engagement and ownership of their learning. Action research makes learning more responsive, relevant, and engaging for students by using personalised strategies, building stronger relationships, encouraging active learning, and giving teachers and students more power. With this method, the classroom is turned into a community of learners who work together to ensure that teaching and learning are constantly changing to meet the needs and potentials of all students (Yosief et al., 2024).

## Literature review

### Student engagement in teaching and learning

Motivation is viewed as both a required component, a prerequisite for student learning engagement, and a necessary element for student engagement in learning. Afzal and Crawford (2022) state that engagement and motivation are needed to improve student learning outcomes. An essential goal of education is to get students excited about learning so that they become motivated to do well and achieve their full potential. Active learning necessitates student participation and interest in the classroom, and students must be encouraged to do so (Tas, 2016). Highly motivated students exert a deliberate effort to engage in classroom activities. Therefore, it is crucial to ascertain students' motivation levels to promote active engagement in class. Students who exhibit higher levels of engagement, possess a stronger sense of ownership and demonstrate a more significant investment in their academic achievements are more inclined to monitor their progress and learning actively. Student engagement in learning is not only an end in itself but also a means to the end of students achieving good academic outcomes (Gao et al., 2022). This is important because authentic engagement may lead to higher academic achievement throughout student life.

Involving students in their learning is an effective way to guarantee that it is meaningful and stimulating. It requires collaboration between the student and the teacher to develop learning and a love of learning. Various issues may impede students' motivation, including mediocre teachers, a flawed curriculum, dislike of a subject, family issues, and instability. "Thoughtful teaching is done by and with students, and not just to them, and students become powerful advocates for their improvement" (Berger et al., 2014, p. 72).

Phrases like student participation, involvement, and engagement are frequently used interchangeably as factors that increase academic performance. Students who are actively involved are an excellent sign of various elementary classroom cultural characteristics, and there is often a significant increase in learning for them. When teacher success is measured as a consequence of student performance, teachers wanting to stay in control may hesitate to focus on improving student ownership of the learning process (Hackett et al., 2021).

Self-monitoring and goal-setting are effective methods for encouraging student ownership because they reduce teachers' time to monitor and evaluate students' progress (Beasy et al., 2022). Teachers would like to focus their efforts on improving student engagement and motivation. Self-assessment allows students to develop transferrable skills that can be used in other areas of learning, such as group projects, teamwork, critical thinking, problem-solving, and leadership roles in the teaching and learning process (Yan & Brown, 2017). It is an effective method for developing standards of knowledge by comparing what they know to the learning objectives and allowing students to improve their work. Students' motivation and engagement often translate into their energy and drive to learn and work well. When the students realise their full potential in school, it significantly influences how much they enjoy and like school (Wang et al., 2022). When students and teachers recognise the importance of the engagement concept and believe that assessments and class activities incorporating this concept are more engaging, it is most probable that students will be motivated to maximise their knowledge and skills. This is what this study attempts to accomplish.

Students can enhance teaching practices and contribute to teachers' professional development. According to the literature (Bartkowiak-Théron et al., 2020), a two-way and multifaceted context helps students feel less alienated and allows for more in-depth learning. Increasing student engagement is frequently viewed as a one-way street, with academics exerting considerable effort to engage students (Clynes, 2009). However, Hood (2012) it argues that when students engage with the subject, the academic's teaching and learning strategies can be drawn from a broader range, improving the learning environment (Hood, 2012).

### **Action research**

Action research approaches are frequently used to improve higher education institutions' curriculum and assessment development strategies. The goal is to ensure that student engagement is a catalyst for learning, emphasising that students must create their understanding rather than having it transmitted to them by teachers (Walton, 2011). According to current knowledge, an assessment can contribute to student involvement (Weaver & Esposto, 2012). It can increase students' participation if they believe the evaluation is fair and appropriate for the subject, especially in group work (Benning, 2022). Utilising the action research methodology to examine students' engagement when there is a change in assessment and class activities will reveal whether the shift has been adequate. Also, the lack of resources and

operational issues are barriers to involving students in their learning (Crabtree et al., 2021). This suggests that service departments and academics must collaborate to ensure that students have access to the resources that will enable them to be more engaged in their studies.

Action research is a problem-solving strategy that entails an iterative data collection, analysis, and reflection process (McTaggart, 1994; Pino-James, 2018). Kurt Lewin coined the term "action research" in 1944 to describe the process of enquiry and investigation that occurs when action is taken to solve a problem. Currently, the term describes a contemplative investigation process carried out to enhance comprehension and application. "Action" pertains to implementing change, while "research" pertains to acquiring enhanced knowledge about learning. In action research, participants examine their educational practice systematically and make informed decisions that can motivate students to become more involved in their education, as done in the case of this research. There are numerous advantages to using action research to investigate teaching and learning. The cyclical nature of action research, which includes planning, action, observation, and reflection, encourages critical review and provides a framework for engagement and improvement (Hodgson et al., 2013).

Educational practitioners and professionals frequently use action research to examine and enhance their pedagogy and practice (Embury et al., 2020). Action research involves the active participation of all stakeholders, bridges the gap between theory and practice, and contributes to generating practical knowledge and implementable results. Using reflective and critical analysis, action research offers a systematic approach to enhancing education and deep learning by placing the practitioner-researcher at the centre (Pennisi et al., 2023).

### **Authentic assessments and quality of engagement**

Teachers can use assessments and class activities to encourage students to take responsibility for their learning (Shepard et al., 2018). Providing students with opportunities to validate their understanding and knowledge while maintaining control over learning objectives through assessment and class activities planning will help foster a culture of student engagement. Furthermore, the types of assessments and learning activities the teachers conduct determine the engaging learning environments for students (Pino-James, 2018). Effective teachers use formative assessments and class activities to fine-tune their responses to individual students' learning needs, identify the sources of student misunderstandings, and determine when students are ready to learn the next step. Authentic assessments use real-world scenarios to engage students in their learning. They make connections, form relationships, and have prior knowledge and skills that allow for multiple solutions due to diverse viewpoints (Moon et al., 2005). Formative assessments lead to various positive outcomes as students take ownership of their learning and see the classroom as a collaboration between themselves and the teachers. Teachers can ensure student engagement and lessen the challenges in evaluation planning by using formal and

informal assessments and in-class activities, student input, and teacher-team collaboration (Gono & de Moraes, 2023). Teachers' assessment practices must change if assessments and classroom activities are to be used to increase student engagement. Teachers should supplement their assessments and class activities with pedagogical alternatives that present novel concepts and skills while engaging students in novel and appropriate learning experiences (Black & Wiliam, 2018). Assessments that elicit higher-order thinking skills from students, such as performance-oriented tasks, increase student engagement in learning and, as a result, help to improve student development.

Teachers are fundamental in creating educational settings that encourage student involvement. Their methods can benefit the quality of engagement and learning.

“Learning environment and individual learning processes cannot be treated as if they exist separately. They both influence one another in a continuous interplay. The influence of instruction will never be direct because of the complexity of mediating variables. In between learning and instruction stand the learner's perceptions of teaching, assessment, course content and structure” (Vermetten et al., 2002, p. 264).

Academics and policymakers have all been interested in improving student learning. Any effort to enhance student engagement must acknowledge the student's uniqueness, and the challenge for educators is determining how this may be accomplished within a group education system. The word “engagement” is often used by educators and others engaged in the scholarship of teaching. “Student engagement represents the time and effort students devote to activities that are empirically linked to desired outcomes of college and what institutions do to induce students to participate in these activities” (Kuh, 2009, p. 683).

Examining the polar opposite of engagement—alienation—is one way to put the notion of engagement into stark perspective. Students' involvement cannot be increased if the teacher focuses only on learning methods (whether deep or superficial) (Mann, 2001). Alienation and disengagement in the student have significant implications beyond just poor learning results; they can result in student withdrawal from the subject and failure to finish the course (Morinaj et al., 2019). An overseas student or a student entering higher education after high school may meet a culture in which the values and views of the participants are likely to be significantly different from their prior experience, posing difficulties for the teachers. Aside from the stress of conforming to academic rigour set by their teachers, students have joined an altogether new and foreign social grouping of similar students, resulting in even more potential for alienation from their classmates.

Establishing more robust connections between the students and lecturers may help increase student involvement (Mann, 2001). It should be stressed that creating a community in the classroom environment may significantly increase student involvement. Some students can achieve high levels of engagement and thought independence with little

assistance. For the others in the cohort, engagement stems from a keen sense of community and acknowledgement. For subjects that require a lot of interaction, such as practical subjects, integration into the class community is highly desirable to encourage students to have a sense of belonging and engagement, resulting in a more fulfilling and enjoyable process than for those without much affiliation (Kember et al., 2001). Some students' primary motivation for tertiary education is to earn a suitable qualification that leads to a decent career after graduation (Bryson & Hand, 2007). For these students, assessments have become the most crucial aspect of learning. However, other students demonstrated genuine enthusiasm for studying on their initiative.

Cultivating robust connections with students is imperative for enhancing student engagement in their educational pursuits (Bartkowiak-Théron et al., 2020). This may present difficulties for some teachers, as their ability to facilitate engagement may be lacking. According to some researchers, encouraging their students to take ownership of their learning may result in less control over the scope and delivery of course material. Empirical research consistently demonstrates that engaging teachers requires facilitation skills and more adaptable lesson plans, which some teachers may lack (Mayeaux & Olivier, 2022; Shea et al., 2015).

This research was undertaken to find pedagogical solutions, as the preceding discussion states that the advantages of engagement outweigh the disadvantages because it helps students develop analytical, evaluation, and collaborative skills. Engagement also promotes creativity and innovation because mindfulness and interaction engage students and empower them to shape their learning experiences. The ultimate goal of engagement is to empower learners to direct their education (Sun & Yang, 2023). As long as a rich learning environment and a strong motivation to learn are present, students will assume the most responsibility for their learning, with the teacher as a facilitator. Students who view a task as meaningful, enjoyable, and challenging become engaged and believe the study is worthwhile. The students' primary concern is that they get it right. They can acquire knowledge at an elevated level, apply it in other contexts, and retain it.

## Methodology

Because action research is characterised by its emphasis on practical consequences and reflective nature, it is well-suited for circumstances in which teachers propose and evaluate instructional innovations. For these reasons, action research was selected for the current study. Action research consists of five phases of inquiry: identification of problem areas, collection and organisation of data, interpretation of data, action based on data, and reflection (Alpert et al., 2022).

Three researchers engaged in action research to explore how they could improve student engagement in one subject of the Bachelor of Culinary Management degree (Kitchen Production and Operations subject – taught by Teacher 1) and two subjects of the Bachelor of Hospitality Management degree (Restaurant Service subject – taught by Teacher 2 and Food and Beverage Knowledge subject – taught by Teacher

3). The three subjects were selected as they are practical-oriented subjects, and this project aimed to improve student engagement, encourage them to take responsibility for their learning in practical training and education settings and effect positive classroom change. Each teacher focused on one subject that they typically teach. Ethics approval for this project was obtained from the institute's Ethics Committee to ensure that ethical and legal research requirements were met. The setting for this research was a renowned institute for hospitality and culinary management in Australia. The research was conducted in the regularly scheduled Bachelor of Culinary Management and Bachelor of Hospitality Management classes over two semesters (12 weeks per semester) with the typically enrolled students and without any specific selection of students or modifications to the regular class situation. This research was carried out in weekly classes of the three practical-oriented subjects over 12 weeks of classes in Semester One. The class length for two subjects was six hours each, and the length of the third subject was four hours. The data was collected by observing the students in the regular practical classes and the quality of the assessments submitted by the students before and after the action research project.

The Kitchen Production and Operations class had nine students, the Restaurant Service class had six students, and the Food and Beverage Knowledge class had 16 students in the first semester. After observing students in the classes of these subjects for the first four weeks of Semester 1, new modified assessments and activities for the three subjects involved were planned as part of the cyclical nature of this action research. The class activities and the assessments were modified by the three teachers consulting together and ensuring that the new class activities and assessments met and adhered to the Intended Learning Outcomes prescribed for each subject. The activities that follow - implementing new activities and assessments for the next eight weeks of the semester, collecting data through ongoing observation of student engagement in the practical classrooms, evaluating the results of the revised assessments and activities, and reflecting on the outcomes by critical review, assisted in establishing a framework for repeating the Semester One process in Semester Two. The information collected for this process was noted in a tabular form with the following headings: the subject's name, activities and assessments prior to action research, modified activities and assessments, outcomes post-action research, and evidence (see Table 1 below). The above process was repeated for another 12 weeks in the next semester with a new batch of enrolled students in the three subjects. Kitchen Production and Operations had 14 students, Restaurant Service had ten students, and Food and Beverage Knowledge had 32 students in the second iteration.

The teachers observed student engagement in practical classes, modified the class activities and assessments, and then re-observed the impact of these changes, implementing a structured approach to data coding and interpretation, ensuring the reliability and validity of the observations. To achieve consistency across observations, all teachers utilised the same criteria for student engagement. This included specific behaviours such as taking the initiative in the classroom, collaborating with other students, possessing

curiosity and inquiry, asking questions, actively participating in class discussions, demonstrating enthusiasm for practical tasks, and contributing to discussion during the class discussions and the de-brief sessions. To maintain a high level of consistency and objectivity in observations, teachers went to each other's classes as observers and moderated assessments of each other's subjects. These checks involved multiple observers (the teachers, in this case) independently coding the same observed events or behaviours according to the predetermined criteria and then comparing their coding to assess agreement. Discrepancies in coding were discussed among the teachers to clarify misunderstandings and refine the application of the coding criteria. This process ensured that all teachers were aligned in their understanding and application of the observation criteria, thereby enhancing the reliability of the observations across different classes and subjects. Such a systematic approach to observing and analysing student engagement allowed for a more objective assessment of the effectiveness of changes to teaching strategies and classroom activities, providing valuable insights into how best to foster an engaging learning environment.

During this research process, iterative steps were employed. The commencement of the enquiry involved the development of an enquiry question: What strategies can be utilised to increase students' level of engagement? The subsequent stage of the research design involved modifying the assessment and activities in three subjects and noting the change in student engagement pre- and post-changes. Throughout the process of data collection, the conduct, social exchanges, and engagement of students within the classroom setting were subject to observation. Furthermore, data collection also includes teachers' discussions with students and their feedback. The process of reflection led to additional modifications in the activities and assessments, incorporating techniques and altered activities and assessments that encouraged heightened engagement. Evaluating and assessing the efficacy of these modifications and interventions required repeated observations of student involvement to verify the sustained increase in engagement. Implementing these changes leads to an enhancement in the overall quality of the learning experience.

The American educational philosopher John Dewey (1997) said: "We do not learn from experience...we learn from reflecting on experience". See Figure 1 below for the conceptual framework.

## **Findings and discussion**

### **Subject 1**

Before action research for the Kitchen Production and Operations class took place, the teacher usually compiled the menu and recipes for the dishes and ordered the ingredients for the class. This activity was modified, and students were asked to design the menus for production and service in the training restaurant, including writing the recipes and ordering the ingredients for the class. The modified activity significantly increased their collaboration and teamwork abilities. They demonstrated initiative and took charge of

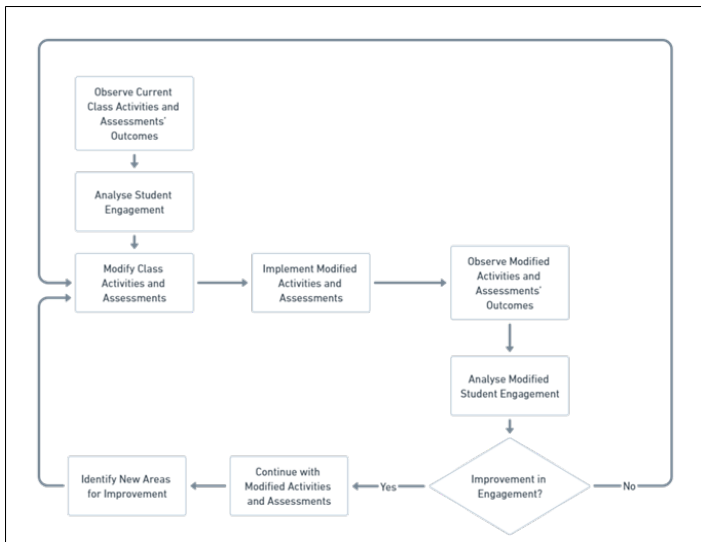


Figure 1: Conceptual framework for action research.

their education when they overcame obstacles or discovered solutions independently, with the teacher as a guide. The programme encouraged and inspired students to assume greater personal responsibility for their Kitchen Production and Operations education. Students could research, try new dishes, and develop original menu concepts (See Table 1 below).

The Kitchen Production and Operations class (Teacher One) monitored the student's progress and helped those who required assistance. After modifying the activity, students were significantly more motivated, interested, and responsible. During after-meal service de-briefing sessions, students were more interested in discussing the positive and negative aspects of the service, and they offered suggestions and comments for improvement. Overall, the student's learning experience was enhanced due to their increased engagement and sense of ownership in their education (Gilmore, 2023). Customers admired the students' efforts, as reflected on the dining customer feedback cards. This suggests that the modified activity improved customer service and learning. In addition, the activity produced a variety of positive outcomes, including increased accountability and responsibility among students, enhanced teamwork and collaboration, observable improvements in motivation, engagement, and self-direction, and positive customer feedback. These results suggest that subject modifications to Kitchen Production and Operations enhanced student learning.

Table 1: Activities and assessments for action research methodology and their outcomes.

Item No.	Name of Subject	Activity/Assessment prior to Action Research	Modified Activity/Assessment	Outcomes Post-Action Research	Evidence
1	Kitchen Production and Operations (Teacher 1)	The teacher compiled the entire menu and the recipes for the menu items for the Kitchen Production and Operations class. The teacher also ordered the ingredients needed for the classes (12 classes in the semester) based on the recipes to be cooked in the class by the students.  The students followed the recipes, prepared them, and served them to the training restaurant guests. There seemed to be a lack of collaboration, exchange of ideas, and discussion between the students in the class on improving the dishes cooked during the day. The students lacked initiative. The teacher conducted the post-	Students were asked to design the menus and write the recipes for the items on the menu to be produced in the kitchen and served in the training restaurant. The students had to order the ingredients needed for the class based on the recipes to be cooked in the class.  This allowed the students to display their creativity, introduce ethnic dishes, create a community, and develop teamwork and ownership. After producing and serving their recipes, prepared them, the students were asked to evaluate the menus they created and identify potential issues in the menu design.	The changed approach in the classroom encouraged/motivated students to take more responsibility for their learning.  Students were autonomous and allowed to explore, experiment and formulate menu ideas – while the teacher observed their learning and guided them.  Students took the initiative in groups to develop menus and recipes and order the ingredients for the recipes for every class. In turns, every student acted as the head chef (prior to action research, the teacher was the head chef)	Teamwork and collaboration improved substantially.  It was noticed that the students actively took a role in leading their learning.  There were noticeable improvements in motivation, engagement and self-direction. The student head chef conducted a post-production/service de-brief. Students were more involved in discussing positive and negative aspects of the service and suggested ideas/comments for improvement. The students took the initiative to create a head chef report to be filled collectively (after de-
				Students took the initiative in groups to develop a plan to solve a task or problem in multiple ways, especially when there was absenteeism in class.	briefing) and given to the teacher at the end of the class.  Positive customer feedback through the customer comment cards increased.  This change in approach for this subject received some good and positive responses from students.
2	Restaurant Service (Teacher 2)	The teacher will plan and manage individual lunch or dinner service at the on-campus Training Restaurant.  This included the teacher allocating three students into a group to act as "Managers" or the Management group for a meal period (one member of the group functioned as the manager and the other two as assistant managers; one assistant manager was responsible for bar and cashiering, and the second was in charge with floor service).  The teacher decided on these duties for the meal period service managers and also assigned duties to the other students or non-managerial students.	With the modification to the approach, the three managers had to decide how to allocate the supervisory duties for restaurant meal period service among themselves. They were responsible for planning a restaurant service and, via a pre-service briefing, advising all staff (other students) on the various duties they had to perform before, during and after the meal service.  The student managers were evaluated and marked collectively on developing and implementing the service plan, providing appropriate feedback, and conducting a de-brief at the end of the service.  Student management teams were assessed on the service management of their assigned meal service. This practical assessment has three (3) components: <ul style="list-style-type: none"><li>• Pre-Service coordination,</li></ul>	Both the management team and the students worked hard to accomplish the mission of providing high-quality service. The management team endeavoured to achieve this objective with the assistance and collaboration of every student, and the instructor standing by to provide direction and advice.  It was observed that students/staff worked with the student managers and were able to come up with restaurant service that was both more creative and flexible. Everyone on the team did what they could to improve the customer experience.  The de-briefing document created a structure for students. Holistically, it gives parameters for others in the class to contribute	Reflection and evaluation of the meal period service were more analytical.  More in-depth post-briefing feedback.  Student feedback on the Management Group's role was more constructive, without fear or favour and driven by the students.  The de-briefing document created a structure for students. They are more motivated to collaborate, communicate, and take the initiative in the de-briefing. This de-briefing document assisted students with fulfilling the requirements of one of the assessments.
		It was noticed that the students would come to the class, wait for the teacher to allocate duties for the meal service, and take no initiative in planning and delivering the food and beverage service.  At the end of the meal service, the teacher conducted a post-meal service de-brief session on how the meal period transpired and what areas needed improvement.	<ul style="list-style-type: none"><li>• Post-service coordination and</li><li>• Peer and customer feedback</li></ul> The student managers for the meal period were asked to lead the post-service meal (this was made part of an assessment, and managers were graded for this), conduct de-brief with the students, and the teacher was present to guide the group. The students introduced a de-briefing document containing topics for the management team to discuss and reflect upon. The students also created this de-briefing document emphasising more excellent reflection and areas where improvements were necessary. (Previously, students utilised the de-brief document template provided by the instructor for practice and evaluation.)	feedback: Debriefings are held to critique, provide feedback, and communicate where improvements must be made to the restaurant's service. They also allow the next group of managers to learn from the previous group's experience.  Previously, de-briefings were rushed and ad hoc in the past due to a lack of specific guidelines. The student-created de-briefing document became a scaffold for the conclusion of service and an excellent tool for assessment and assessment grading.	The assessment became richer in content and reflected an honest perspective regarding their progress in creating a restaurant service experience for the guests.



Item No.	Name of Subject	Activity/Assessment prior to Action Research	Modified Activity/Assessment	Outcomes Research	Post-Action Research	Evidence
3	Food and Beverage Knowledge (Teacher 3)	Food and wines are served in this class. As an assessment, the students wrote reflections for every food and wine-tasting session. This formed an assessment in itself. Based on the food and wine tasted in the tutorial, students must write a concise, highly relevant and detailed understanding of the theoretical concepts related to the sensory evaluation of food and beverage commodities and matching principles. Appropriate and consistent tasting terminology/vocabulary is to be used in writing this reflection.	The teacher invited a peer to the class to observe the teacher-student interaction. The peer advised Teacher 3 that to increase student participation, the teacher should consider dividing the class into small groups, providing more time, and waiting longer for student responses/discussion to food and wine pairing questions. Previously, the questions asked by the teacher were rapid-fire. Allowing students to process their thoughts on the wines and food in pairs or putting students in small groups would help them participate more fully in discussions and answer questions in class as a team.	Taking the comments of the peer partner into consideration, students were divided into smaller groups (groups of four students each) and asked to discuss the pairing of food and beverage within their groups; each group then presented their pairing perspectives—this increased student participation in their discussion. The primary objective was for students to independently initiate and sustain an academic debate. Multiple iterations of student-led discussions effectively boost student engagement, intellectual investment, and participation in academic discussions. Our lesson plans and evaluations must support our primary objective, which is for these discussions to lead to increased autonomy in		The changes in approach to class participation resulted in improved written output for the assessment based on the tutorial discussion on food and wine pairing. Students were more confident in their ability to pair food and wine, more engaged in class discussion, and had a deeper understanding of the topic. The quality of the written reflection also improved. After the change, the reflection was more evident, concise, highly relevant, and detailed in its understanding of the theoretical concepts related to the sensory evaluation of food and beverage commodities and matching principles.
		It was noticed that only a few students participated in food and wine pairing discussions in the tutorial, and most students were not confident, as evidenced by their unwillingness to answer the questions or participate in class discussions and the quality of the written reflection assessment on tasting.		sustaining classroom conversations as part of daily practice so that students can use these skills and dispositions to become actively engaged, productive participants and take ownership of their learning.		There was an appropriate and consistent use of food and wine/beverage terminology/vocabulary throughout.

## Subject 2

In the Restaurant Service class (Teacher 2), groups of three students acting as 'Managers' or the Management group ran meal period service in the training restaurant (one assumed the manager and the other two as assistant managers). One assistant manager was responsible for bar and cashing, and the second was for floor service. The teacher in the earlier semesters decided on these duties for the three managers of the meal period service. With the modification to the approach, the three managers had to determine how to allocate the responsibilities for restaurant meal period service among themselves. They were responsible for planning a restaurant service and appropriately communicating via verbal and written instruction to all other students acting as staff in the training restaurant. The three managers allocated duties to students acting as staff to do the necessary mise-en-place, conduct a briefing before service delivery, and provide feedback at the end of the meal service period. The teacher evaluated and marked the student managers collectively in developing and implementing the service plan, providing appropriate feedback to the students acting as staff in the training restaurant, and conducting a debrief at the end of the service.

Both the management team and the students worked hard to accomplish the mission of providing high-quality service. The student managers or the management group team endeavoured to achieve this objective with the assistance and collaboration of every student/staff member, with the teacher standing by to provide direction and advice. It was observed that students/staff worked with the student managers and were able to come up with restaurant service that was both more creative and flexible. Everyone on the team did what they could to improve the customer experience (see Table 1 above).

## Subject 3

In the subject of Food and Beverage Knowledge class (Teacher 3), students were provided food and wines. It was noticed that only a few students participated in food and wine pairing discussions, and most students were not confident, as evidenced by their unwillingness to answer the questions or participate in class discussions on food and wine pairing. The teacher invited a peer to the class to observe the teacher-student interaction (Johnston et al., 2022). The peer advised the teacher to increase student participation by providing more time and waiting longer for student responses/discussion to food and wine pairing questions. Taking the comments of the peer partner into consideration, students were divided into smaller groups (groups of four students each) and asked to discuss the pairing of food and beverage within their groups; each group then presented their pairing perspectives. All these changes increased student participation in their discussion.

The primary objective was for students to independently initiate and sustain an academic debate. Multiple iterations of student-led discussions effectively boosted student engagement, intellectual investment, and participation in academic discussions. The revised lesson plans and evaluations must support the primary objective, which is for these discussions to lead to increased autonomy in sustaining classroom conversations as part of daily practice so that students can use these skills and dispositions to become actively engaged, productive participants and take ownership of their learning. Allowing students to process their thoughts on the wines and food in pairs or putting students in small groups would help them participate more fully in discussions and answer questions in class as a team. Students independently investigated and evaluated the food and beverage pairing through peer-led discussions. As a result, the students exhibited an enhanced capacity to comprehend the basic principles of sensory evaluation and matching. Providing increased autonomy in the classroom enabled students to take charge of their learning. Implementing the modified tutorial activity yielded a noteworthy enhancement in the overall calibre of the written reflections on food and wine pairing. The students clearly and concisely understood the subject matter directly applicable to the task. The responses provided exhibited a comprehensive understanding of the concepts, as evidenced by the consistent and precise application of tasting terms and related vocabulary. The revised evaluation and activity facilitated the process of composing reflective essays for students on food and wine pairing, enriched their comprehension of the subject matter, and bolstered their engagement, ultimately leading to better academic achievements. Based on the available evidence, the instructional style and material alterations positively impacted the students' understanding and engagement with food and wine tasting (see Table 1 above).

## Further discussion

The three examples above highlight that in assisting students to take ownership of their learning, teachers can help them achieve academic and behavioural goals. On

the other hand, empowering students to take ownership of their education can help them complete their education more quickly, teach them essential skills such as setting and attaining goals, and help them develop their independence. To take control of their own learning experiences, students must clearly understand the intended learning objectives (Macfarlane & Tomlinson, 2017). Student ownership is based on specific learning objectives since learning objectives inform students about what is expected of them regarding performance. Clearly defined learning objectives, class activities and assessments are powerful tools for increasing student ownership of learning (Reeve et al., 2020).

John Dewey, the father of reflection, defined reflection as "Active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it and the further conclusion to which it tends" (Dewey, 1997, p. 118).

In the case of this research, reflective practice enabled teachers to change class activities and assessments to improve their teaching and better meet their students' learning needs. This improvement occurs as the reflective practice enhances teaching, learning, and comprehension, an essential component of teachers' professional development (Pineda et al., 2022). When a teacher conducts rigorous self-examination, they better understand themselves, their practices, and their students. The authors incorporated reflective practice into their action research to increase awareness of their underlying beliefs and assumptions about learning and teaching and modify their lesson delivery.

Assessments and activities were modified to make them challenging for the entire population of students being evaluated. They are precise regarding what is expected of students and related to the learning objectives or specific learning outcomes. Teachers provided students with various classroom participation opportunities because they understood the importance of meeting each student's diverse needs and learning requirements (Beasy et al., 2022). The significance of the new activities motivated and engaged the students, who appeared engaged and believed the modified learning activity worthwhile.

Most assessment activities and modifications led to collaborative learning, a potent engagement booster for learning activities. When students collaborate to create performance on the assessment and class activities, they better understand the standards taught and the assessed performance outcomes (Qureshi et al., 2023). Students who collaborated effectively during this research paid more attention, primarily because they felt more connected to their peers during activities. Teachers promoted motivation and cooperative learning by avoiding homogeneous groups, grouping students based on their abilities, encouraging individual accountability through the distribution of various roles, and evaluating individual and group performance.

Teachers must address increasing student achievement while modifying instruction, classroom assessments and activities, and curriculum assessments. Encouraging students to take responsibility for their learning as a fundamental component of formative teaching techniques is one strategy

for increasing student achievement. By taking an active role in the process, students can be taught how to make decisions about their learning. When students become personally invested in their education, they understand their learning objectives, gather and document evidence of their learning, and assess and set new learning objectives, assisting them in achieving the ultimate goal of increasing their academic success (Beasy et al., 2022). Because they are overburdened with other responsibilities, teachers may overlook opportunities to increase student ownership of learning. When teachers assist their students in taking an active role in their education by providing opportunities for them to do so, they contribute to the student achievement goals.

Activity and assessment modifications challenged students and made them work harder. The difficulty stimulated higher-order thinking, which requires students to delve deeper into problems and ideas. Examples of higher-order thinking included making judgments about information, manipulating data on assessments and activities to reveal deeper meanings, analysing various facets of an idea, combining data from multiple sources to produce new interpretations, and applying concepts and theories to novel situations.

By creating engaging, open-ended tasks and assessments that focus on real-world skills and allow students to choose how to measure and reflect on their progress, students can be empowered to take ownership of their learning. Researchers must figure out what motivates and engages students. Many advocates and researchers advocate for educational institutions to include students in educational decision-making to increase student engagement and investment in themselves and their education (Kurtz et al., 2019). Minimise subject barriers by promoting students' integration of their existing knowledge from other subjects with the new information they acquire in their current subject. Educational technology, such as Moodle (Learning Management System), can help students develop skills to participate in 21st-century learning. Active learning strategies such as discussion, debate, group work, and team projects require students to deepen their understanding of the content.

The implementation of student-centred strategies has yielded numerous favourable consequences. At the outset, there has been a notable augmentation in the array of prospects for pedagogy and acquisition. Educators can enhance student engagement, comprehension, and retention by tailoring instructional approaches to accommodate individual students' unique needs and interests. The second aspect of development pertains to the heightened involvement of students in the decision-making process. This allows students to actively participate in voicing their opinions regarding matters that directly impact their educational experiences. When students engage in activities that cultivate a sense of ownership and responsibility, they are empowered to take control of their educational trajectory.

Furthermore, the active engagement of students has facilitated enhanced communication and collaboration among their peers. Engagement in collaborative endeavours

and undertakings fosters cooperation, facilitates the exchange of ideas, and cultivates essential interpersonal skills among students. Students' increasing autonomy and intrinsic motivation have facilitated heightened activity in their educational pursuits. When students possess a heightened sense of ownership and autonomy in their educational endeavours, they tend to be more inclined towards actively pursuing knowledge, establishing academic objectives, and demonstrating perseverance. Implementing these improvements has resulted in a more efficient and inclusive educational setting.

## Conclusion

Action research is a systematic enquiry approach that seeks to produce novel knowledge with practical applications, enhance current methodologies, and enhance students' teaching and learning experiences. Together with motivation, engagement is viewed in the literature as very important for enhanced learning outcomes for all students (Raza et al., 2020). Through action research, teachers can continue to improve their teaching methods and students' learning outcomes. Additionally, they can contribute to educational knowledge. It encourages educators to make decisions with more deliberation and evidence. This allows them to tailor their instructional strategies to the needs of their students.

This action research study aimed to determine how to engage students more in their education and give them a sense of ownership over their education. This research shed light on strategies for engaging students in learning and giving them greater control over their education. Enhancing student engagement within the classroom has positively improved motivation, engagement, and comprehension levels.

Implementing student-centred teaching strategies, such as project-based learning and collaborative activities, has significantly increased student learning engagement. Participating students displayed tremendous enthusiasm, motivation, and active engagement in their learning experiences.

The findings also demonstrated the importance of fostering a supportive learning environment in the classroom that values students' ideas and encourages them to act independently. When students chose what they would learn, established goals, and evaluated what they had learned, they felt greater responsibility and ownership over their work. Students had a stronger sense of agency and were more invested in learning when their opinions were considered during lesson planning and grading.

The findings of this action research study increased our understanding of engaging students in their learning and empowering them to take ownership of their educational experiences. Employing student-centred practices, creating a welcoming classroom environment, and valuing student voice and choice enable educators to equip students with the skills necessary to be active participants in their learning. This study provides teachers and researchers with vital information they can use to assist all students in learning in

a meaningful and long-lasting way.

The action research process has assisted the researchers/teachers in understanding what is going on in the classroom and identifying changes that will improve teaching and learning. Action research is a valuable tool for answering how to enhance the effectiveness of educational strategies and student performance. This action research has generated qualitative data that can be used to improve the subject matter being taught and the instructional methods used in the curriculum. Teachers can make informed decisions with the help of action research.

## Implications for research and practice

Employing action research to evaluate tactics for increasing student engagement is vital for creating a more effective learning environment (Gibbs et al., 2004). Teachers can enhance student engagement and foster student ownership of learning by implementing the action research methodology in their classrooms through collaborative strategies, reflective practice, individual interventions, and continuous improvement. Both teachers and students can use action research to create engaging and meaningful learning experiences.

Students and teachers who participate in action research acquire research skills through their participation in research. The abilities acquired include critical thinking, problem-solving, data collection, reflection, and analysis. These skills are transferable and can be applied in different contexts. In addition, students learn self-directed learning, self-evaluation and self-regulation. With the help of action research, educators can continue enhancing the content of their lessons and the knowledge their students acquire.

Action research encourages educators to critically evaluate their teaching practices and make modifications based on added information. Teachers can find effective ways to engage students in learning and encourage them to assume responsibility for their education by participating in the research process. The study's findings have demonstrated that the implementation of student-centred teaching strategies, such as project-based learning and group activities, has resulted in a significant increase in students' level of interest in learning. This study has confirmed a positive correlation between increased student engagement and student achievement and that student engagement is a significant factor in academic performance.

With action research, educators can tailor their interventions to each student's needs and circumstances. If they continuously plan, act, observe, and reflect on their actions, teachers can determine what is most effective for their students. Then, based on their research findings, they can adjust their interventions.

## Limitations and future research

Typically, the researcher for action research is the teacher or practitioner. Data collection and analysis can become subjective and biased. Measuring student ownership of their learning and engagement is challenging, as these concepts are subjective actions and perceptions. The three teachers attended each other's classes as observers and used the same student engagement criteria to overcome subjectivity and bias and ensure consistency across observations. These checks involved the three teachers coding the same events or behaviours independently according to predetermined criteria and then comparing their coding to assess agreement. Teachers discussed coding discrepancies to clarify and improve coding criteria. This process aligned teachers' understanding and application of observation criteria, improving observation reliability across classes and subjects. A systematic approach to observing and analysing student engagement allowed for a more objective assessment of teaching strategy and classroom activity changes, revealing how to create an engaging learning environment.

Nonetheless, it is essential to comprehend students' perspectives, and these developments bode well for future research in the field. Future research can examine grades from assessments (such as exams, activities, and projects) from previous years when the course was taught in the conventional format to compare grades before and after the changes. Action research usually requires a substantial investment of resources and time. The process includes planning, executing, collecting data, analysing, and evaluating the results. Finding a balance between regular teaching duties and action research participation may be demanding for teachers. Stakeholders such as students and co-workers may oppose implementing changes based on action research. Resistance can be caused by numerous factors, such as a lack of understanding of the new strategy, a reluctance to change routines or a lack of confidence in the efficacy of action research. To overcome opposition, teachers must effectively collaborate and communicate with others.

## Acknowledgement

We thank Associate Professor Melanie Williams – Associate Dean (Scholarship) – William Angliss Institute, for guiding us with this research project and reviewing this research document. We are very grateful to have had her assistance.

## References

Abdullah, J., Mohd-Isa, W. N., & Samsudin, M. A. (2019). Virtual reality to improve group work skill and self-directed learning in problem-based learning narratives. *Virtual Reality*, 23(4), 461-471. <https://doi.org/10.1007/s10055-019-00381-1>

Afzal, F., & Crawford, L. (2022). Students' perception of engagement in online project management education and its impact on performance: The mediating role of self-motivation. *Project Leadership and Society*, 3, 100057. <https://doi.org/10.1007/s10055-022-00057-1>

[doi.org/https://doi.org/10.1016/j.plas.2022.100057](https://doi.org/10.1016/j.plas.2022.100057)

Alpert, S., Dean, M., & Ewell, W. (2022). Preparing educational leaders through action research. *Action Research*, 21(1), 124-143. <https://doi.org/10.1177/14767503221143872>

Bartkowiak-Théron, I., McShane, A. L. J., & Knight, M. G. (2020). Departing from anonymous and quantitative student feedback: Fostering learning and teaching development through student evaluations. *Journal of Applied Learning and Teaching*, 3(Sp. Iss. 1), 118-128. <https://doi.org/10.37074/jalt.2020.3.s1.16>

Beasy, K., Morrison, R., Coleman, B., & Mainsbridge, C. (2022). Reflections of a student engagement program designed and delivered by academics. *Journal of Applied Learning & Teaching*, 5(1), 40-51. <https://doi.org/10.37074/jalt.2022.5.1.7>

Benning, T. M. (2022). Reducing free-riding in group projects in line with students' preferences: Does it matter if there is more at stake? *Active Learning in Higher Education*, 14697874221118864. <https://doi.org/10.1177/14697874221118864>

Berger, R., Rugen, L., & Woodfin, L. (2014). *Leaders of their own learning: Transforming schools through student-engaged assessment*. San Francisco, CA: John Wiley and Sons Inc.

Black, P., & William, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551-575. <https://doi.org/10.1080/0969594X.2018.1441807>

Bryson, C., & Hand, L. (2007). The role of engagement in inspiring teaching and learning. *Innovations in Education and Teaching International*, 44(4), 349-362. <https://doi.org/10.1080/14703290701602748>

Chan, P. E., Graham-Day, K. J., Ressa, V. A., Peters, M. T., & Konrad, M. (2014). Beyond involvement: Promoting student ownership of learning in classrooms. *Intervention in School and Clinic*, 50(2), 105-113. <https://doi.org/10.1177/1053451214536039>

Clynes, M. P. (2009). A novice teacher's reflections on lecturing as a teaching strategy: Covering the content or uncovering the meaning. *Nurse Education in Practice*, 9(1), 22-27. <https://doi.org/10.1016/j.nepr.2008.03.007>

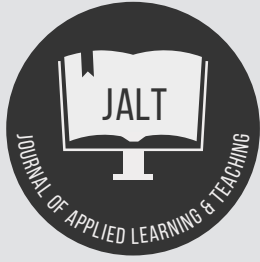
Crabtree, R. M., Briggs, P., & Woratschek, H. (2021). Student engagement and barriers to implementation: The view of professional and academic staff. *Perspectives: Policy and Practice in Higher Education*, 25(4), 144-150. <https://doi.org/10.1080/13603108.2021.1946446>

Dewey, J. (1997). *How we think*. Dover Publications. <https://archive.org/details/howwethinkrestat00dewerich>

Embury, D. C., Parenti, M., & Childers-McKee, C. (2020). A charge to educational action researchers. *Action Research*, 18(2), 127-135. <https://doi.org/10.1177/1476750320919189>

- Gao, X., Guo, F., & Coates, H. (2022). Contributions to the field of student engagement and success. *Higher Education Research & Development, 41*(1), 62-74. <https://doi.org/10.1080/07294360.2021.2008326>
- Gibbs, P., Angelides, P., & Michaelides, P. (2004). Preliminary thoughts on a praxis of higher education teaching. *Teaching in Higher Education, 9*(2), 183-194. <https://doi.org/10.1080/1356251042000195367>
- Gilboy, M. B., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior, 47*(1), 109-114. <https://doi.org/https://doi.org/10.1016/j.jneb.2014.08.008>
- Gilmore, H. (2023). The (academic) road less travelled: From dropout to recovery. *Journal of Applied Learning and Teaching, 6*(2), 402-413. <https://doi.org/10.37074/jalt.2023.6.2.14>
- Gono, S., & de Moraes, A. J. (2023). Student appraisals of collaborative team teaching: A quest for student engagement. *Journal of Applied Learning and Teaching, 6*(1), 222-233. <https://doi.org/10.37074/jalt.2023.6.1.26>
- Hackett, J., Kruzich, J., Goulter, A., & Battista, M. (2021). Tearing down the invisible walls: Designing, implementing, and theorizing psychologically safer co-teaching for inclusion. *Journal of Educational Change, 22*(1), 103-130. <https://doi.org/10.1007/s10833-020-09401-3>
- Hodgson, Y., Benson, R., & Brack, C. (2013). Using action research to improve student engagement in a peer-assisted learning programme. *Educational Action Research, 21*(3), 359-375. <https://doi.org/10.1080/09650792.2013.813399>
- Hood, A. (2012). Whose responsibility is it? Encouraging student engagement in the learning process. *Music Education Research, 14*(4), 457-478. <https://doi.org/10.1080/14613808.2012.703174>
- Johnston, A. L., Baik, C., & Chester, A. (2022). Peer review of teaching in Australian higher education: A systematic review. *Higher Education Research & Development, 41*(2), 390-404. <https://doi.org/10.1080/07294360.2020.1845124>
- Kamali, J. (2023). Metamorphosis of a teacher educator: A journey towards a more critical self. *Journal of Applied Learning & Teaching, 6*(2), 252-259. <https://doi.org/10.37074/jalt.2023.6.2.8>
- Kember, D., Lee, K., & Li, N. (2001). Cultivating a sense of belonging in part-time students. *International Journal of Lifelong Education, 20*(4), 326-341. <https://doi.org/10.1080/02601370117754>
- Kuh, G. D. (2009). The national survey of student engagement: Conceptual and empirical foundations. *New Directions for Institutional Research, 2009*(141). <https://doi.org/10.1002/ir.283>
- Kurtz, J. B., Lourie, M. A., Holman, E. E., Grob, K. L., & Monrad, S. U. (2019). Creating assessments as an active learning strategy: What are students' perceptions? A mixed methods study. *Medical Education Online, 24*(1), 1630239. <https://doi.org/10.1080/10872981.2019.1630239>
- Liu, F., Wang, X., & Izadpanah, S. (2023). The comparison of the efficiency of the lecture method and flipped classroom instruction method on EFL students' academic passion and responsibility. *SAGE Open, 13*(2), 21582440231174355. <https://doi.org/10.1177/21582440231174355>
- Macfarlane, B., & Tomlinson, M. (2017). Critiques of student engagement. *Higher Education Policy, 30*(1), 5-21. <https://doi.org/10.1057/s41307-016-0027-3>
- Mann, S. J. (2001). Alternative perspectives on the student experience: Alienation and engagement. *Studies in Higher Education, 26*(1), 7-19. <https://doi.org/10.1080/03075070020030689>
- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal, 37*(1), 153-184. <https://doi.org/10.3102/00028312037001153>
- Mayeaux, A. S., & Olivier, D. F. (2022). Professional kinship using social media tools: Bridging and bonding to develop teacher expertise. *Journal of Applied Learning and Teaching, 5*(Sp. Iss. 1), 27-34. <https://doi.org/10.37074/jalt.2022.5.s1.3>
- McTaggart, R. (1994). Participatory action research: Issues in theory and practice. *Educational Action Research, 2*(3), 313-337. <https://doi.org/10.1080/0965079940020302>
- Moon, T. R., Brighton, C. M., Callahan, C. M., & Robinson, A. (2005). Development of authentic assessments for the middle school classroom. *Journal of Secondary Gifted Education, 16*(2-3), 119-133. <https://doi.org/10.4219/jsge-2005-477>
- Morinajpaj, J., Marcin, K., & Hascher, T. (2019). School alienation and its association with student learning and social behavior in challenging times. In E. N. Gonisa, & M. S. Lemos (Eds.), *Motivation in education at a time of global change* (Volume 20, pp. 205-224). Emerald Publishing Limited. <https://doi.org/10.1108/S0749-742320190000020010>
- Pennisi, S., Lathrop, A., & Pilato, K. A. (2023). How can practitioner action research support the design, implementation, and evaluation of on-campus mental health and addiction services? *Educational Action Research, 1*-19. <https://doi.org/10.1080/09650792.2023.2229402>
- Pineda, J. L. d. L., Villanueva, R. L. d. D., & Tolentino, J. A. M. (2022). Virtual focus group discussions: The new normal way to promote reflective practice. *Reflective Practice, 23*(2), 190-202. <https://doi.org/10.1080/14623943.2021.2001322>
- Pino-James, N. (2018). Evaluation of a pedagogical model for student engagement in learning activities. *Educational Action Research, 26*(3), 456-479. <https://doi.org/10.1080/09650792.2017.1354771>
- Pudjiarti, E. S., Rini, W., & Wae, D. (2023). Cooperative

- Learning: An effective approach for improving student engagement and achievement. *Educational Administration: Theory and Practice*, 29(2). <https://doi.org/10.17762/kuey.v29i2.685>
- Qureshi, M. A., Khaskheli, A., Qureshi, J. A., Raza, S. A., & Yousufi, S. Q. (2023). Factors affecting students' learning performance through collaborative learning and engagement. *Interactive Learning Environments*, 31(4), 2371-2391. <https://doi.org/10.1080/10494820.2021.1884886>
- Ramsden, P. (2003). *Learning to teach in higher education*. Routledge. <https://doi.org/10.4324/9780203507711>
- Raza, S. A., Qazi, W., & Umer, B. (2020). Examining the impact of case-based learning on student engagement, learning motivation and learning performance among university students. *Journal of Applied Research in Higher Education*, 12(3), 517-533. <https://doi.org/10.1108/JARHE-05-2019-0105>
- Reeve, J., Cheon, S. H., & Jang, H. (2020). How and why students make academic progress: Reconceptualizing the student engagement construct to increase its explanatory power. *Contemporary Educational Psychology*, 62, 101899. <https://doi.org/https://doi.org/10.1016/j.cedpsych.2020.101899>
- Shea, J., Joaquin, M. E., & Gorzycki, M. (2015). Hybrid course design: Promoting student engagement and success. *Journal of Public Affairs Education*, 21(4), 539-556. <https://doi.org/10.1080/15236803.2015.12002219>
- Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice*, 37(1), 21-34. <https://doi.org/https://doi.org/10.1111/emip.12189>
- Sun, Z., & Yang, Y. (2023). The mediating role of learner empowerment in the relationship between the community of inquiry and online learning outcomes. *The Internet and Higher Education*, 58, 100911. <https://doi.org/https://doi.org/10.1016/j.iheduc.2023.100911>
- Tas, Y. (2016). The contribution of perceived classroom learning environment and motivation to student engagement in science. *European Journal of Psychology of Education*, 31(4), 557-577. <https://doi.org/10.1007/s10212-016-0303-z>
- Vermetten, Y. J., Vermunt, J. D., & Lodewijks, H. G. (2002). Powerful learning environments? How university students differ in their response to instructional measures. *Learning and Instruction*, 12(3), 263-284. [https://doi.org/https://doi.org/10.1016/S0959-4752\(01\)00013-5](https://doi.org/https://doi.org/10.1016/S0959-4752(01)00013-5)
- Walton, J. (2011). A living theory approach to teaching in higher education. *Educational Action Research*, 19(4), 567-578. <https://doi.org/10.1080/09650792.2011.625718>
- Wang, J., Tigelaar, D. E. H., Luo, J., & Admiraal, W. (2022). Teacher beliefs, classroom process quality, and student engagement in the smart classroom learning environment: A multilevel analysis. *Computers & Education*, 183, 104501. <https://doi.org/https://doi.org/10.1016/j.compedu.2022.104501>
- Weaver, D., & Esposto, A. (2012). Peer assessment as a method of improving student engagement. *Assessment & Evaluation in Higher Education*, 37(7), 805-816. <https://doi.org/10.1080/02602938.2011.576309>
- Wheaton, M. (2021). Why student engagement in the accounting classroom matters. *Journal of Applied Learning and Teaching*, 4(1), 72-81. <https://doi.org/10.37074/jalt.2021.4.1.3>
- Yan, Z., & Brown, G. T. L. (2017). A cyclical self-assessment process: Towards a model of how students engage in self-assessment. *Assessment & Evaluation in Higher Education*, 42(8), 1247-1262. <https://doi.org/10.1080/02602938.2016.1260091>
- Yosief, A., Sulieman, M.-S., & Biede, T. (2024). Improving the practices of teacher educators through collaborative action research: Challenges and hopes. *Educational Action Research*, 32(2), 204-221. <https://doi.org/10.1080/09650792.2022.2066147>



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Evaluation of a research training workshop for academic staff in tertiary institutions: A Kirkpatrick model approach

Caroline Ochuko

A

*Faculty of Education, University of Delta Agbor, Delta State, Nigeria*

Alordiah<sup>A</sup>

### Keywords

Academic staff;  
Kirkpatrick model;  
publication strategies;  
research skills;  
research training workshop;  
research visibility;  
tertiary institutions.

### Abstract

Quality research has a positive impact on the development of a country. Literature has shown that there is a limited quantity of quality African research articles in reputable journals. Training and workshops have consistently been shown to have a positive impact on the productivity of academic staff and researchers. If academic staff are trained in writing and publishing research articles, it will go a long way to improve the contribution of African research to reputable journals. Hence, the focus of this study is to evaluate the workshop conducted on writing and publishing academic papers in highly reputable journals for academic staff in tertiary institutions. The study's population consists of academic staff in tertiary institutions in Delta State, Nigeria. The sample comprised 59 academic staff who participated voluntarily in the workshop. Through a pre-questionnaire, a post-questionnaire, a participant satisfaction level questionnaire, and interviews, the four stages of the Kirkpatrick Model were used to assess the effectiveness of the workshop.

### Correspondence

[caroline.alordiah@unidel.edu.ng](mailto:caroline.alordiah@unidel.edu.ng)<sup>A</sup>

### Article Info

Received 30 December 2023  
Received in revised form 17 January 2024  
Accepted 12 March 2024  
Available online 8 April 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.32>

High levels of participant satisfaction and notable advancements in academic writing, publication, and research exposure were observed as a result of the workshop. As they established online research profiles, identified predatory journals, and improved paper preparation, participants actively were able to apply their newly gained abilities. A 12-month post-workshop evaluation found remarkable results, including 18 participants publishing articles in journals with a Scopus index and many more creating profiles on Google Scholar, ResearchGate, and Academic.edu. The study highlights the significance of customised workshops in advancing research abilities and academic recognition by demonstrating a favourable association between customised workshops and increased research capabilities. Future evaluations can use the evaluation model as a useful framework, allowing for well-informed judgments about institutional and educational improvements.

## Introduction

Utilising modern scientific techniques to conduct research appropriately is a sign of quality research. Being prominent in the academic community and having publications in respectable journals are crucial requirements for receiving individual and institutional recognition (Alordiah et al., 2023; McGrail et al., 2006). Many accreditation bodies or organisations need academic personnel to remain engaged in their field, and publishing is a prominent and well-liked means to do so (Northcentral University, 2020; Tella & Onyancha, 2020). Academic publication is the process through which someone or a group makes intellectual content accessible to a general audience. Before findings are published and made publicly accessible, the research process must be followed (Owan & Asuquo, 2022). Skills (introduction, methods, findings, discussion, and references), evaluation (peer review), sharing (publishing of the scholarly materials), and preservation of the content are all necessary for the production of scholarly materials (databases and repositories). A journal is a piece of academic writing published by an accredited publisher, faculty, department, or university. It should have an editorial board and local, national, or worldwide readers, as the situation may be (Tella, 2015). However, a journal with a global reputation can help people and organisations gain more international renown and respect.

Many people think Nigeria's higher education institutions have fallen far behind in innovative research (Alordiah et al., 2023; Alordiah et al., 2021; Can et al., 2018). Two-thirds of articles published in predatory journals are reportedly written in Asia and Africa, primarily in Nigeria, India, Turkey, and Pakistan (Demir, 2018). Consequences of the development of these situations in Nigeria include low research quality, researchers and institutions having a poor standing, a lack of confidence in Nigerian research, and difficulties obtaining research funds (Briggs & Weathers, 2016). African researchers are however encouraged to publish in prestigious publications, including in Web of Science and Scopus (Owan et al., 2023; Alordiah et al., 2021).

The number of articles published in "high-impact" journals (journals indexed in Scopus and Web of Science) and the h-index thresholds are currently used by several tertiary institutions in Nigeria to assess promotion prospects. The management of these institutions argued that by making this move, they would deter academic staff from writing in predatory journals and encourage them to increase visibility for both their institutions and themselves (Owan & Asuquo, 2022). This decision implies that all academic staff members must have accounts on Academia.edu, Research Gate, and Google Scholar. Additionally, they are required to produce works in journals with high impact routinely and have an ORCID number. The reality is that for academic staff at higher institutions, writing for scientific publications remains a challenge and a significant difficulty (Tella & Onyancha, 2020; Habibie, 2015). When academics are putting up papers for publications that are indexed by Scopus or Web of Science (WoS), these difficulties become more glaring. To produce scholarly publications, academics need to have knowledge of how to conduct a literature search, establishing a suitable methodology, conducting an

analysis of the data, and presenting the findings coherently. Academic staff members must also have the ability to write well and present their findings and arguments clearly before submitting to credible publications. Additionally, they require knowledge of how to publicise their articles to a wider audience (Inee et al., 2018).

When it comes to publishing their findings in reputed international peer-reviewed journals, researchers in underdeveloped nations frequently lack the necessary scientific writing abilities. These scholars are not well represented in these publications, and a rising number of their works appear in predatory and dubious journals.. African researchers have a little opportunity for mentorship and inadequate training in writing research articles (Alordiah et al., 2021; Shoko et al., 2021; Sumathipala et al., 2004). Access to university writing centres and research writing training is more prevalent in developed countries than in poorer nations (Shoko et al., 2021; Sumathipala et al., 2004). In most of Nigeria's higher institutions, preparing articles for publication in highly reputable journals is a relatively recent development. As such, it becomes important to plan training on academic writing and research visibility.

In September 2022, a research workshop was held to assist participants who had little to no prior experience writing for highly regarded journals, had never published an article in one, had only recently begun the process of preparing an article for one, or were hoping to increase the visibility of their research articles on a global scale. The Carvimial Research Advisory (CRAD) group planned the workshop. The workshop's main objectives were publishing in respected journals, increasing research visibility, and writing research articles. The workshop had the following sections: introduction, methodology, results, discussion, and conclusion. It is good to conduct such workshops, but the most crucial thing is determining their impact.

One of the fundamental components of any programme or educational process is evaluation. A workshop or programme's objectives are evaluated to ascertain if they have been met. Information is gathered to assess the workshop's effectiveness (Abdulghani et al., 2014; Musal et al., 2008). In addition to ensuring that each trainee has achieved their educational goals, workshop organisers guarantee the programme's overall quality (Abdulghani et al., 2014; Durning et al., 2007). To evaluate academic programmes, several evaluation models have been put forth. But for many years, the Donald Kirkpatrick model (Kirkpatrick & Kirkpatrick, 2006) has been the main organisational design for the assessments of training (Abdulghani et al., 2014; Smidt et al., 2009).

One of the complete methods for assessing training and workshops is Kirkpatrick's evaluation model. This model includes four fundamental evaluation stages, and each level influences the next. The first level is concerned with how the participants perceive the training programme. It gauges participant satisfaction and gathers data on their experiences with the training they received (Kirkpatrick & Kirkpatrick, 2006). The participants will most likely learn new things at Level 2, where they will modify their attitudes and behaviours. If the evaluation of the training programme



involves finding out what information and abilities have been acquired by the participant, the assessors might utilise a pre-questionnaire and post-questionnaire (Abdulghani et al., 2014). The third level assesses if the newfound information, abilities, and attitudes have been applied to the workplace to reflect improvements in conduct and productivity. The participants' enhanced performance results are measured at the fourth level. Kirkpatrick's model is a trustworthy evaluation model. It is valid, reliable, and reasonably cheap. Additionally, evaluators may use both qualitative and quantitative methods and subjective and objective indicators. In affluent nations, the use of the Kirkpatrick Model for reviews of educational initiatives that alter workplace dynamics are consistently reported (AlFaris et al., 2015; Abdulghani et al., 2014; Bailey & Hewison, 2014; Christopher & Young, 2015; Scott et al., 2013; Smidt et al., 2009). However, there are not many such reports from Nigeria.

This study aims to close this gap by using the Kirkpatrick Model to assess the workshop's effectiveness in terms of participant satisfaction, enhancement of their relevant conceptual knowledge and cognitive skills, participant behavioural changes, and the primary outcomes in the form of publications and visibility. As far as we know, this is one of the few studies to evaluate a training workshop in Nigeria for writing and publishing papers in respectable journals. This evaluation method would offer valuable guidance for how workshop developers can evaluate their workshop to determine the impact of their training on participants. It will encourage tertiary institutions to use this workshop on article writing and publications in their various institutions to promote quality research among academic staff.

### Research questions

1. To what extent do participants find the training satisfactory? (Level 1 of Kirkpatrick model)
2. To what degree do the participants acquire the basic knowledge and cognitive skills taught during the training sessions? (Level 2 of the Kirkpatrick model)
3. To what extent did the participants apply what they learned during the training? (Level 3 of Kirkpatrick model)
4. What target outcomes occurred as a result of the training? (Level 4 of the Kirkpatrick model)

### Hypothesis

There is no significant difference between the participants' pre-questionnaire and post-questionnaire scores on the basic knowledge and cognitive skills taught during the training sessions.

### Method

In this evaluation study, an explanatory sequential mixed method (QUAN-qual) is used. A quantitative study was conducted to measure the first and second levels of the Kirkpatrick model. The third and fourth levels of the Kirkpatrick model were measured through qualitative research. The concept of explanatory sequential mixed methods, also known as QUAN-qual, refers to a research design that amalgamates both quantitative and qualitative components systematically. Generally, the quantitative aspect precedes the qualitative part in this design. Utilising an explanatory sequential mixed methods (QUAN-qual) approach within the scope of this study allows for the integration of both quantitative and qualitative elements throughout the research process, in a particular order (Liem, 2018).

### Quantitative study

#### Participants

The study population comprises the academic staff of tertiary institutions in Delta State. Flyers and posters (e-copies and hardcopies) were sent to academic staff, who were encouraged to participate in the workshop. The sample comprised 59 academic staff who volunteered to participate in the workshop. However, only 33 of them completed the pre-questionnaire and post-questionnaire questionnaires. Of these 33 academic staff, 26 (79%) and 7 (21%) were female and male, respectively. Based on the area of specialisation, 13 (40%), 3 (9%), 4 (11%), and 13 (40%) were from the faculties of sciences/engineering, arts, social sciences, and education, respectively. About 25 (77%) have not published any article in a Scopus or Web of Science journal. However, 4 (11%), 1 (3%), 1 (3%), and 1 (3%) of the participants have published 1, 2, 3, 4, and 6 articles in journals indexed in Scopus and Web of Science databases.

#### Materials

A pre-questionnaire titled "Writing, publication and visibility of academic papers in reputable journals A" and a post-questionnaire titled "Writing, publication and visibility of academic papers in reputable journals B" were used. They were developed based on the workshop's objectives and content. Both questionnaires contained the same items. Section One measured the personal information of the participants. Section Two measured the knowledge and skills needed to write articles publishable in reputable journals and had ten items. The third section measured the knowledge and skills required to publish articles in reputable journals. It contains eight items. Section Four measured the knowledge and skills needed to make your articles visible internationally. There were ten items in this section. A 7-point Likert scale was used to measure the items in Sections 2, 3, and 4. Such that Not at all, A little, A little below average, Moderately, A little above average, Well, and Very well were awarded 1, 2, 3, 4, 5, 6, and 7 points, respectively. The items in the pre-questionnaire and post-questionnaire were identical. Experts in measurement and evaluation validated these items. The

reliability of Sections 2, 3, and 4 using Cronbach Alpha were .71, .71, and .73, respectively. The overall reliability coefficient for the full scale was .72. The pre-questionnaire and post-questionnaire were used to measure Level 2 of the Kirkpatrick model. A third questionnaire titled "Participants' satisfaction level" was used to collect data to measure Level 1 of Kirkpatrick's model. This questionnaire has nine items with a 7-point Likert scale. Experts in measurement and evaluation also validated it, and it yielded a reliability coefficient of .76 using Cronbach Alpha.

## Procedures

The pre-questionnaire was administered to the participants before the treatment (training sessions). The training sessions lasted for two days, covering about 14 hours. The topics covered were on writing a good research article 1 (introduction, literature review, discussion, and conclusion sections), Writing a good research article 2 (method and result sections), Writing a good title and abstract, Publishing in highly reputable journals, and having research visibility. After the end of the training, the post-questionnaire and the third questionnaire (Participants' satisfaction level) were administered to the participants. The participants were free to decide if they wanted to complete the questionnaire.

## Analysis

The data was checked to satisfy the assumptions needed for the specific analysis. The pre-questionnaire and post-questionnaire data contained outliers but were both normally distributed. The "Participants' satisfaction level" data did not have outliers but was not normally distributed. This information guided the researcher in determining the right statistical tool to use. Frequency count, percentage, and graphical illustration were used to answer Research Questions One and Two. The Wilcoxon Signed Rank test was used to test the hypothesis at a .05 level of significance.

## Qualitative study

After the end of the workshop, the researcher contacted the participants via e-mail, over the phone, WhatsApp, face-to-face communication or observation to find out their research activities after the workshop. Over 30 of the participants responded to the calls and messages. The researcher asked them how they used the knowledge and skills they learned during the workshop. Also, the researcher observed those who were accessible to determine their present attitude in their place of work concerning the knowledge and skills they acquired during the workshop. The information gathered was used to answer Research Question Three. For Research Question Four, the researcher communicated with them about 12 months after the workshop to find out their current level of visibility and the articles they were able to publish. The information was verified by going to the various websites to confirm their reports. The qualitative data was analysed through content analysis.

## Research question one

To what extent do participants find the training satisfactory?

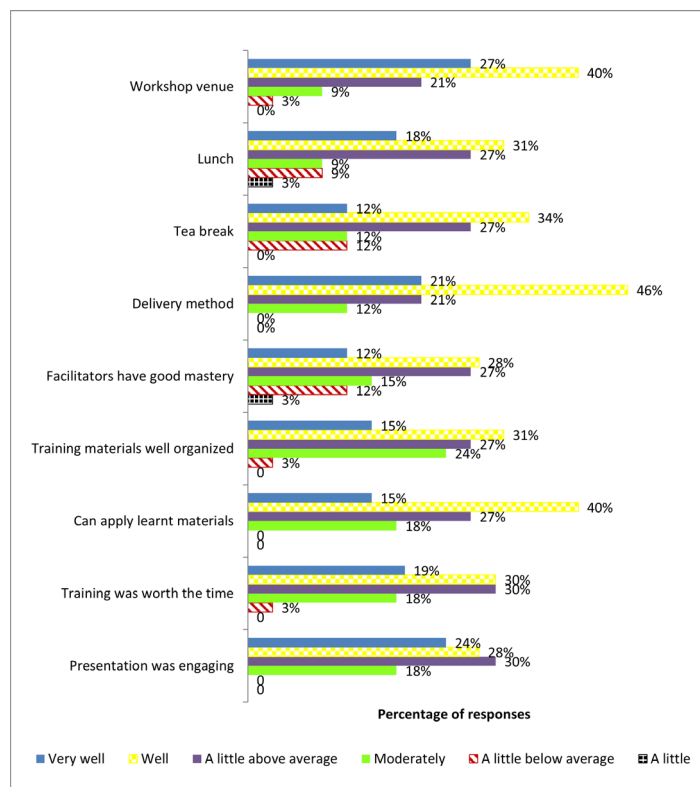


Figure 1: Extent of participants' satisfaction.

The majority of the lecturers were satisfied with the level of engagement of the presentation (very well=24%, well=28%, and a little above average=28%); the training was worth the time (very well=19%, well=30%, and a little above average=30%); and the materials learnt can be applied to their research (very well=15%, well=40%, and a little above average=27%). Many of the lecturers were satisfied with the level at which the training materials were well organised (very well=15%, well=31%, a little above average=27%); facilitators' mastery of the subject matter (very well=12%, well=28%, a little above average=27%); and the delivery method (very well=21%, well=46%, a little above average=21%). A large number of the lectures were satisfied with the tea break (very well=12%, well=34%; a little above average=27%); lunch (very well=18%, well=31%, a little above average=27%); and the workshop venue (very well=27%, well=40%, a little above average=21%).

## Research question two

To what degree do the participants acquire the basic knowledge and cognitive skills taught during the training?

## Writing academic papers

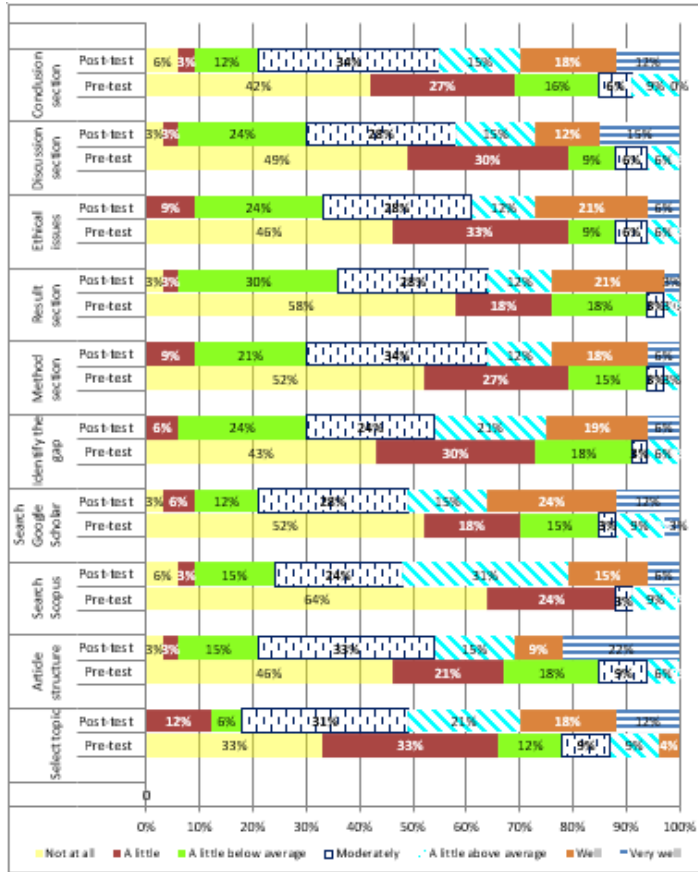


Figure 2: Participants' level of acquisition of knowledge and skills for writing academic papers.

## Publishing in reputable journals

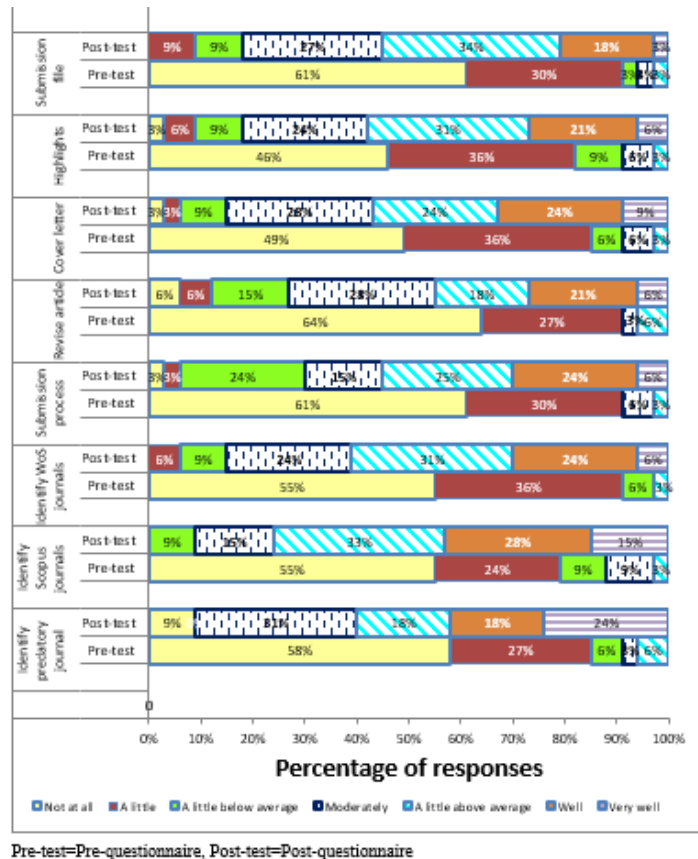


Figure 3: Participants' level of acquisition of knowledge and skills for publishing in highly reputable journals.

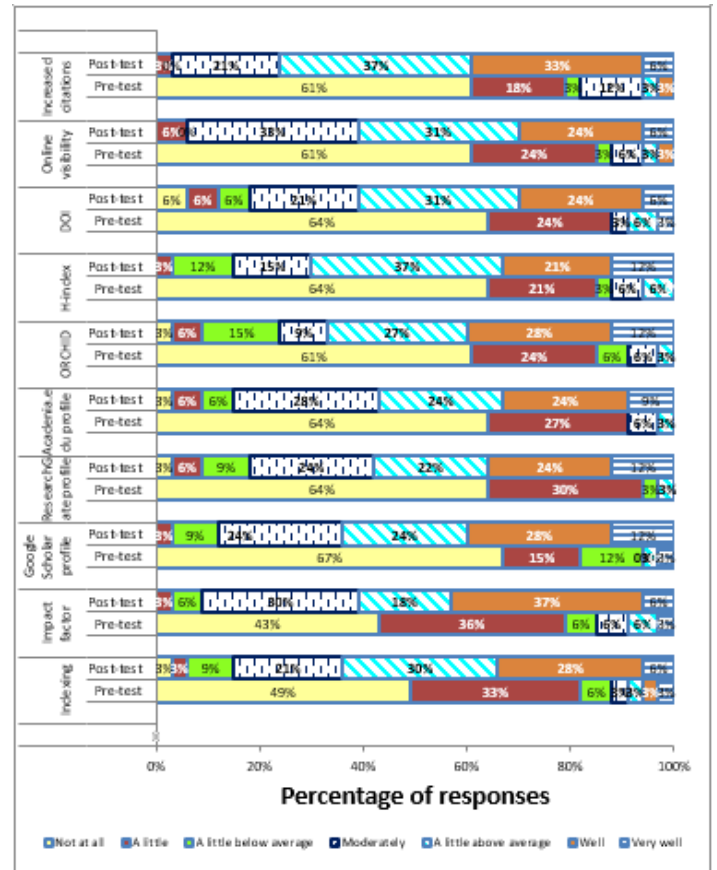


Figure 4: Participants' level of acquisition of knowledge and skills for research visibility.

## Writing academic papers

From the pre-questionnaire, more than 60% of the workshop participants had little or no knowledge of how to select a good topic, structure an article, search for articles in Scopus and Google Scholar databases, identify the gap in a study, write a good methodology and result sections, handle ethical issues, and write the discussion and conclusion sections. This percentage was reduced to 13% in the post-questionnaire. Hence, based on the lecturers' pre-questionnaire and post-questionnaire scores, the lecturers gained more basic knowledge and skills in writing academic papers for reputable journals at the end of the workshop (Figure 2).

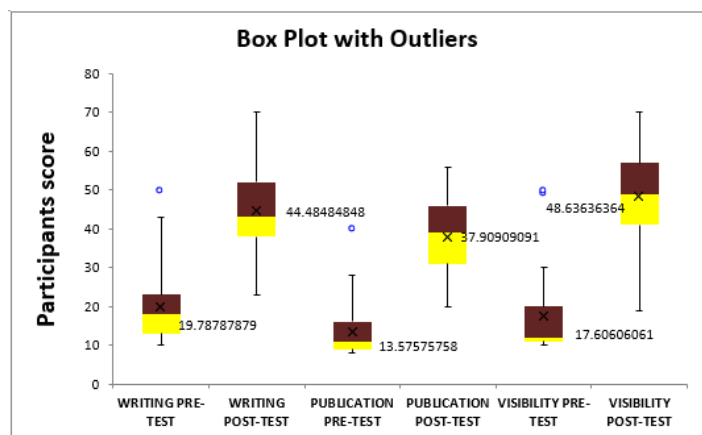
## Publishing in reputable journals

From the pre-questionnaire, more than 70% of the workshop participants had little or no knowledge of how to identify predatory, Scopus, and WoS journals, understand the submission process for reputable journals, revise an article, write cover letters and highlights for reputable journals, and prepare the submission files needed by reputable journals. This percentage was reduced to 13% in the post-questionnaire. Hence, based on the lecturers' pre-questionnaire and post-questionnaire scores, the lecturers gained more basic knowledge and skills in publishing in reputable journals at the end of the workshop (Figure 3).

## Research visibility

From the pre-questionnaire, more than 75% of the workshop participants had little or no knowledge of journal indexing and impact factor, Google Scholar, ResearchGate, Academia.edu profiles, ORCID, DOI, making articles visible online, and strategies to attract more citations. This percentage was reduced to 12% in the post-questionnaire. Hence, based on the lecturers' pre-questionnaire and post-questionnaire scores, the lecturers gained more basic knowledge and skills in research visibility at the end of the workshop (Figure 4).

The box plot in Figure 5 showed that the mean scores for the pre-questionnaire for writing academic articles (19.79), publishing in reputable journals (13.50), and research visibility (17.61) were lower than the mean scores for the post-questionnaire for writing academic article (44.48), publishing in reputable journals (37.91), and research visibility (48.64).



Pre-test=Pre-questionnaire, Post-test=Post-questionnaire

Figure 5: Box plot of participants' pre-questionnaire and post-questionnaire scores.

Table 1: Descriptive statistics for writing, publication, and visibility of articles.

	N	Percentiles		
		25 <sup>th</sup>	50 <sup>th</sup> (Median)	75 <sup>th</sup>
Writing Post-questionnaire	33	37.50	43.00	53.00
Writing Pre-questionnaire	33	12.50	18.00	23.50
Publishing Post-questionnaire	33	30.50	39.00	47.00
Publishing Pre-questionnaire	33	8.50	11.00	16.00
Visibility Post-questionnaire	33	40.50	49.00	57.50
Visibility Pre-questionnaire	33	10.50	12.00	20.50

Table 1 shows the descriptive statistics of pre-questionnaire and post-questionnaire scores. The pre-questionnaire's median (50th percentile) score for writing an article is 18.00, publishing in reputable journals is 11.00, and research visibility is 12.00. While the median score of the post-questionnaire for writing an article is 43.00, publishing in reputable journals is 39.00, and research visibility is 49.00. The difference between these pre-questionnaire and post-questionnaire scores is quite big. The p-values of the Wilcoxon

Table 2: Wilcoxon signed rank test for writing, publication, and visibility of articles.

	N	Mean Rank	Sum of Ranks	Z	Sig
WRITING POST-QUESTIONNAIRE - WRITING PRE-QUESTIONNAIRE	Negative Ranks	0 <sup>a</sup>	.00		5.016 .000
	Positive Ranks	33 <sup>b</sup>	17.00	561.00	
	Ties	0 <sup>c</sup>			
	Total	33			
PUBLISHING POST-QUESTIONNAIRE - PUBLISHING PRE-QUESTIONNAIRE	Negative Ranks	0 <sup>a</sup>	.00		5.014 .000
	Positive Ranks	33 <sup>d</sup>	17.00	561.00	
	Ties	0 <sup>e</sup>			
	Total	33			
VISIBILITY POST-QUESTIONNAIRE - VISIBILITY PRE-QUESTIONNAIRE	Negative Ranks	0 <sup>a</sup>	.00		5.015 .000
	Positive Ranks	33 <sup>b</sup>	17.00	561.00	
	Ties	0 <sup>c</sup>			
	Total	33			

a. WRITING POST-QUESTIONNAIRE < WRITING PRE-QUESTIONNAIRE  
 b. WRITING POST-QUESTIONNAIRE = WRITING PRE-QUESTIONNAIRE  
 c. WRITING POST-QUESTIONNAIRE > WRITING PRE-QUESTIONNAIRE  
 d. PUBLISHING POST-QUESTIONNAIRE < PUBLISHING PRE-QUESTIONNAIRE  
 e. PUBLISHING POST-QUESTIONNAIRE = PUBLISHING PRE-QUESTIONNAIRE  
 f. PUBLISHING POST-QUESTIONNAIRE > PUBLISHING PRE-QUESTIONNAIRE  
 g. VISIBILITY POST-QUESTIONNAIRE < VISIBILITY PRE-QUESTIONNAIRE  
 h. VISIBILITY POST-QUESTIONNAIRE = VISIBILITY PRE-QUESTIONNAIRE  
 i. VISIBILITY POST-QUESTIONNAIRE > VISIBILITY PRE-QUESTIONNAIRE

Signed Rank test are 0.000 for writing the article, publishing in reputable journals, and research visibility. These p-values indicate that the pre-questionnaire and post-questionnaire scores (medians) differ significantly. Therefore, the training has significantly improved the participants' knowledge and skills in writing articles, publishing in reputable journals, and research visibility because the median of the post-questionnaire scores is significantly higher than the pre-questionnaire scores (Table 2).

## Research question three

To what extent did the participants apply what they learned during the training?

The post-workshop follow-up revealed that many participants have started preparing articles to be published in Scopus or Web of Science journals. Here are some of the participants' comments in Figure 6:

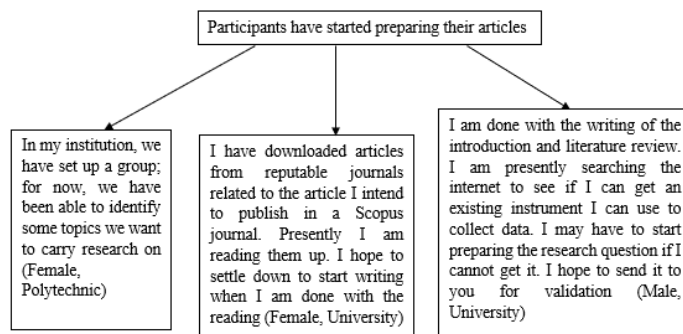


Figure 6: Participants' preparation for article writing after the workshop.

The majority of the participants can now verify whether a journal is predatory or not. A female College of Education lecturer has this to say:

"O! I feel so bad that I wasted so many resources publishing in predatory journals. I can now identify a predatory journal. In recent times I have taken

my time to look at the composition of the editorial board and the time between when you submit and when your paper will be accepted for publication. I also check whether the journal is indexed in Scopus, DOAJ, AJOL, and other reputable databases. I cannot be fooled again.”

The participants also said they had started the process of opening their Academic.com, ResearchGate, and Google Scholar accounts. A sample of the participant responses is shown in Figure 7.

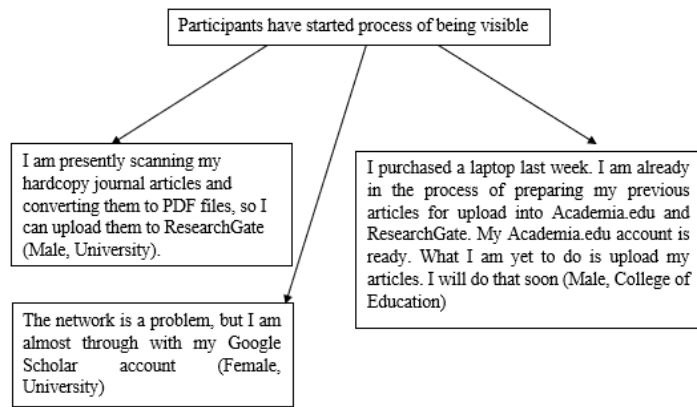


Figure 7: Participants' preparation for research visibility after the workshop.

#### Research question four

What targeted outcomes occur as a result of the training?

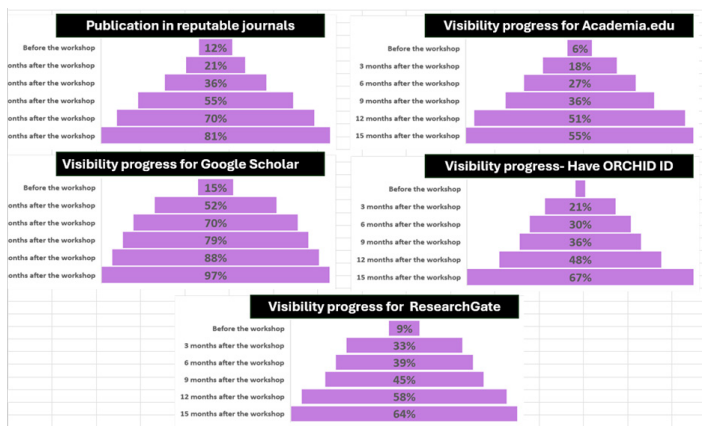


Figure 8: Impact metrics of post-research training workshop.

The primary objective of this segment is to decipher the target outcomes resulting from the research training workshop, employing Level 4 of the Kirkpatrick Model, which emphasises the measurement of results. The data sets collected at various intervals post-workshop shed light on multiple facets: publication in reputable journals, visibility progress on platforms like Google Scholar, ResearchGate, Academia.edu, and the acquisition of an ORCID ID.

#### Publication in reputable journals

The tabulated data delineates a significant escalation in the number and percentage of participants publishing in reputable journals post-workshop. Initially, a mere 4 participants, constituting 12%, were engaged in such scholarly activities. However, as time progressed, there was a conspicuous amplification: 15 months post-training witnessed a commendable 81% of participants (n=27) making contributions to reputable journals (Figure 8). Such a trajectory underscores the efficacy of the workshop in fostering academic productivity and dissemination of scholarly output.

#### Visibility progress for Google Scholar

The progression of visibility on Google Scholar manifests a palpable upward trend. Prior to the workshop, a modest 15% (n=5) were visible on this platform. Yet, post-training intervals reveal an exponential surge, culminating in an impressive 97% (n=32) visibility 15 months post-workshop (Figure 8). This remarkable augmentation accentuates the workshop's instrumental role in enhancing academic visibility, thereby augmenting the dissemination and impact of research endeavours.

#### Visibility progress for ResearchGate

Analogous to the trends observed on Google Scholar, visibility on ResearchGate witnessed a robust ascent after the workshop. A mere 9% (n=3) of participants exhibited visibility on this platform pre-workshop. However, the ensuing months showcased a consistent elevation, with 15 months post-training registering a commendable 64% (n=21) visibility (Figure 8). This trajectory corroborates the workshop's efficacy in augmenting the digital footprint and scholarly interaction of participants on ResearchGate.

#### Visibility progress for Academia.edu

The visibility progression on Academia.edu mirrored the trends observed across other platforms, albeit with slight variations. Initial pre-workshop engagement stood at a modest 6% (n=2). Nevertheless, subsequent evaluations portrayed a steady ascent, culminating in 55% (n=18) visibility 15 months post-training (Figure 8). Such a trajectory underscores the workshop's efficacy in fostering an active scholarly presence on diverse digital platforms, thereby enhancing academic collaboration and dissemination.

#### Visibility progress - possession of ORCID ID

The acquisition of an ORCID ID, a pivotal identifier for academic recognition, exhibited a consistent and substantial growth trajectory post-workshop. Merely 3% (n=1) of participants possessed an ORCID ID pre-workshop. However, subsequent intervals showcased a remarkable augmentation, with 67% (n=22) acquiring this essential identifier 15 months post-training (Figure 8). This trajectory underscores the

workshop's efficacy in equipping participants with essential tools for academic recognition and collaboration.

## Discussion

The study results revealed positive findings for the writing and publishing of articles in reputable journal workshops at all four levels of the Kirkpatrick model. The present workshop programme was valuable in contributing to the writing and publication of articles in reputable journals and the visibility of the participants. The research outcomes have shown that the participants were willing to produce articles to be published in reputable journals. Previous studies have shown that such hand-on-training programmes also support learners to improve their research (Fuller et al., 2005) and reinforce changes at the institutional level (Abdulghani et al., 2014; Nestel et al., 2004).

Kirkpatrick's first level of evaluation assesses participants' reactions to the facilitators' delivery method, the training materials, the workshop venue, lunch, and learning activities. The majority of the participants were satisfied with the workshop. This result was in line with that of Abdulghani et al. (2014). Many institutions used the first level of the Kirkpatrick model as the sole means of workshop evaluation (Abdulghani et al., 2014). Positive satisfaction does not ensure learning and subsequent application of the workshop content. Kirkpatrick's second level evaluates the extent of learning among the participants (Kirkpatrick & Kirkpatrick, 2006). The participants' basic knowledge and cognitive skills were high for writing the academic paper, publishing in reputable journals, and research visibility. The post-questionnaire scores were significantly higher than the pre-questionnaire scores for writing academic papers, publishing in reputable journals, and research visibility. This improvement may be attributed to the quality delivery of the training materials by the experienced resource specialists and the interactive sections that followed each training session. This improvement is similar to findings reported in previous studies (Abdulghani et al., 2014).

The evaluation of Kirkpatrick's third and fourth levels also showed marked improvement in the participants' article writing, publishing, and research visibility skills and knowledge. The data reveals a notable transformation from passive consumers of knowledge to active contributors to the academic discourse and a heightened visibility on Google Scholar, ResearchGate, and Academic.com platforms. Also, there was an exponential growth in ORCID ID acquisition. This transformation is indicative of the workshop's success in not only imparting theoretical knowledge but also inculcating practical skills essential for navigating the intricate landscape of academic publishing. It signifies that participants' research endeavours are not confined to the boundaries of their respective institutions but resonate across global academic communities, fostering interdisciplinary collaboration and knowledge exchange. It also implies that the workshop's curriculum, pedagogical approaches, and resources resonated with participants' evolving academic aspirations and professional paths. Research training is an important strategy for improving researchers' skills, productivity, and visibility (Vinnikova et

al., 2021; Devlin & Radloff, 2014; Francis et al., 2009).

The multifaceted outcomes elucidated herein accentuate the workshop's transformative impact on participants' academic trajectories, institutional profiles, and scholarly activities. However, future endeavours may necessitate additional longitudinal evaluations that may elucidate sustained impacts, emergent trends, and evolving academic landscapes, thereby informing iterative refinements, strategic interventions, and transformative initiatives within tertiary institutions.

The findings showed that the participants could use the knowledge and skills acquired during the workshop and yielded results. If this improvement was recorded just 15 months after the workshop, there is hope that there may be a more remarkable improvement in the next 3 years. Several studies evaluating workshops, training, and programmes using Kirkpatrick's model did not effectively measure the third and fourth levels (Cahapay, 2021; Dewi & Kartowagiran, 2018; Aryadoust, 2017; Steele et al., 2016). This study has shown results for the 3rd and 4th levels. This is an additional improvement to what some previous works have found.

However, there are some limitations associated with this study. The sample size was small. The reason for this was that both the workshop attendance and filling in of the pre-questionnaire and post-questionnaire were voluntary. Another limitation was that the researcher did not receive feedback through the institutions because the participants were drawn from several institutions. It is easier to collect information for the fourth level of Kirkpatrick's model if the participants of the workshop were from the same institution. In addition, observation of the activities and behaviours of the participants was limited to only the institution to which the researcher had access. Other factors, like the researchers' previous knowledge, may have contributed to the findings in Level Four of Kirkpatrick's model. Nevertheless, the knowledge and skills learnt during the workshop have significantly contributed to the study's results in the fourth level of Kirkpatrick's model.

The researcher believes that the findings of this study could significantly influence the higher education publication landscape in Nigeria, as well as potentially benefit other regions in Africa and beyond. By showcasing the success of tailored training programmes in improving academic staff members' research skills and knowledge of reputable publishing practices, this study encourages the implementation and refinement of similar initiatives across different institutions. These efforts can lead to enhanced overall research output and international collaboration opportunities, ultimately elevating the global standing of participating universities and countries.

Possible future research directions include conducting longitudinal studies on the lasting impacts of targeted workshops, investigating variation in effectiveness due to factors such as discipline, career stage, gender, and location, examining challenges in scaling up personalised workshops, comparing various research skill development techniques, studying industrial collaborations for real-world application of research, focusing on capacity building for

early-career researchers from underrepresented groups, and incorporating technology like text prediction software, AI-driven plagiarism detection, and VR simulations for immersive data analysis in research training.

## Conclusion

Participants' feedback is important and useful for improving and conducting academic workshops. In addition, this workshop was found to be effective and meet the needs of the participants who can improve institutional research capacity and visibility. There is a need to re-evaluate the fourth level after six months to get a true picture of the impact of this workshop on the participants. Furthermore, academic workshops, training, and programmes should be evaluated using evaluation models like Kirkpatrick's model.

## Implications of the findings

### Theoretical implications:

This study carries several theoretical implications that contribute to the understanding of research training workshops and their impact on academic staff in tertiary institutions.

1. The study showcases the effectiveness of Kirkpatrick's four-level evaluation model in comprehensively assessing the outcomes of a research training workshop.
2. This study extends the understanding of how workshops influence participants' learning and subsequent actions.
3. The study emphasises the importance of participant feedback for designing effective workshops.
4. The high participant satisfaction levels highlight the workshop's success in meeting participant expectations.
5. This understanding can guide future workshop organisers in tailoring their content and delivery methods to align with participants' needs and preferences.

### Empirical implications:

Empirically, this study has practical implications for workshop organisers, academic staff, and institutions:

1. The findings offer valuable insights for designing and delivering effective research training workshops.
2. Workshop organisers can take cues from the participants' reported satisfaction and improved knowledge to refine their training content, methods, and delivery, ultimately enhancing the workshop's impact.

3. The study showcases the potential of workshops to empower academic staff in underdeveloped regions with vital skills for successful research endeavours.
4. Participants' improved skills in academic writing, publishing, and research visibility demonstrate the practical benefits of targeted training initiatives.
5. Institutions can leverage the study's outcomes to enhance their research capacity and visibility.
6. The 12-month follow-up assessment offers a novel approach to gauging the long-term impact of workshops.
7. Future evaluations can adopt similar longitudinal approaches to gain a deeper understanding of how workshops influence participants over time.

## Recommendations

Based on the findings and implications of the study, several recommendations are suggested to enhance the effectiveness of research training workshops for academic staff in tertiary institutions:

1. Workshop organisers should ensure that the content is well-aligned with the specific needs and challenges faced by academic staff in the given context.
2. Customising the training content to address participants' existing knowledge gaps and research requirements can enhance the relevance and impact of the workshop.
3. Incorporating interactive learning methods, such as hands-on exercises, case studies, and group discussions, active engagement and deeper understanding among participants can be facilitated.
4. Implementing a long-term follow-up mechanism, similar to the 12-month assessment in this study, is recommended.
5. Institutions should recognise the value of research training workshops and actively support their staff's participation.
6. Ensuring that the workshop is facilitated by experienced and knowledgeable trainers is crucial. Continuous assessment and refinement of workshop content and methods are essential.
7. Collaborating with academic institutions can enhance the workshop's reach and impact.

## References

Abdulghani, H. M., Shaik, S. A., Khamis, N., Al-Drees, A. A., Irshad, M., Khalil, M. S., ... & Isnani, A. (2014). Research methodology workshops evaluation using the Kirkpatrick's model: Translating theory into practice. *Medical Teacher*,

36(Supp. 1), S24-S29. <https://doi.org/10.3109/0142159X.2014.886012>

AlFaris, E., Naeem, N., Irfan, F., Qureshi, R., Saad, H., Al Sadhan, R. E., ... & Van der Vleuten, C. (2015). A one-day dental faculty workshop in writing multiple-choice questions: An impact evaluation. *Journal of Dental Education*, 79(11), 1305-1313. <https://doi.org/10.1002/j.0022-0337.2015.79.11.tb06026.x>

Alordiah, C. O., Osagiede, M. A., Omumu, F. C., Okokoyo, I. E., Emiko-Agbajor, H. T., Chenube, O., & Oji, J. (2023). Awareness, knowledge, and utilisation of online digital tools for literature review in educational research. *Heliyon*, 9(1), e12669. <https://doi.org/10.1016/j.heliyon.2022.e12669>

Alordiah, C. O., Owamah, H. I., Ogbinaka, E. J. A., & Alordiah, M. O. (2021). Nigeria's low contribution to recognized world research literature: Causes and remedies. *Accountability in Research*, 28(8), 471-491. <https://doi.org/10.1080/08989621.2020.1855984>

Aryadoust, V. (2017). Adapting levels 1 and 2 of Kirkpatrick's model of training evaluation to examine the effectiveness of a tertiary-level writing course. *Pedagogies: An International Journal*, 12(2), 151-179. <https://doi.org/10.1080/1554480X.2016.1242426>

Association of College & Research Libraries (ACRL). (2003). *Principles and strategies for the reform of scholarly communication*. College and Research Libraries News. <http://www.ala.org/acrl/publications/whitepapers/principlesstrategies>

Association of College & Research Libraries (ACRL). (2020). *Principles and strategies for the reform of scholarly communication 1*. <http://www.ala.org/acrl/publications/whitepapers/principlesstrategies>

Bailey, C., & Hewison, A. (2014). The impact of a 'Critical Moments' workshop on undergraduate nursing students' attitudes to caring for patients at the end of life: An evaluation. *Journal of Clinical Nursing*, 23(23-24), 3555-3563. <https://doi.org/10.1111/jocn.12642>

Briggs, R. C., & Weathers, S. (2016). Gender and location in African politics scholarship: The other white man's burden?. *African Affairs*, 115(460), 466-489. <https://doi.org/10.1093/afraf/adw009>

Cahapay, M. (2021). Kirkpatrick model: Its limitations as used in higher education evaluation. *International Journal of Assessment Tools in Education*, 8(1), 135-144. <https://doi.org/10.21449/ijate.856143>

Can, N., Tursunbadalov, S., & Keles, I. (2018). An assessment of scientific research in Nigerian universities. *Journal of Economics and Social Research*, 5(10). [https://www.researchgate.net/publication/338913963\\_An\\_Assessment\\_of\\_Scientific\\_Research\\_in\\_Nigerian\\_Universities](https://www.researchgate.net/publication/338913963_An_Assessment_of_Scientific_Research_in_Nigerian_Universities)

Christopher, M. M., & Young, K. M. (2015). Awareness of "predatory" open-access journals among prospective veterinary and medical authors attending scientific writing

workshops. *Frontiers in Veterinary Science*, 2, 22. <https://doi.org/10.3389/fvets.2015.00022>

Demir, S. B. (2018). Predatory journals: Who publishes in them and why?. *Journal of Informetrics*, 12(4), 1296-1311. <https://doi.org/10.1016/j.joi.2018.10.008>

Devlin, M., & Radloff, A. (2014). A structured writing programme for staff: Facilitating knowledge, skills, confidence and publishing outcomes. *Journal of Further and Higher Education*, 38(2), 230-248. <https://doi.org/10.1080/0309877x.2012.722194>

Dewi, L. R., & Kartowagiran, B. (2018). An evaluation of internship program by using Kirkpatrick evaluation model. *Research and Evaluation in Education*, 4(2), 155-163. <https://doi.org/10.21831/reid.v4i2.22495>

Durning, S. J., Hemmer, P., & Pangaro, L. N. (2007). The structure of program evaluation: An approach for evaluating a course, clerkship, or components of a residency or fellowship training program. *Teaching and Learning in Medicine*, 19(3), 308-318. <https://doi.org/10.1080/10401330701366796>

Francis, K., Mills, J., Chapman, Y. B., & Birks, M. (2009). Doctoral dissertations by publication: Building scholarly capacity whilst advancing new knowledge in the discipline of nursing. *International Journal of Doctoral Studies*, 4, 97-106. <https://doi.org/10.28945/695>

Fuller, A., Hodkinson, H., Hodkinson, P., & Unwin, L. (2005). Learning as peripheral participation in communities of practice: A reassessment of key concepts in workplace learning. *British Educational Research Journal*, 31(1), 49-68. <https://doi.org/10.1080/0141192052000310029>

Habibie, P. (2015). An investigation into writing for scholarly publication by novice scholars: Practices of Canadian anglophone doctoral students. *Electronic Thesis and Dissertation Repository*, 3281. <https://ir.lib.uwo.ca/etd/3281>

Kirkpatrick, D., & Kirkpatrick, J. (2006). *Evaluating training programs: The four levels*. Berrett-Koehler Publishers. [http://dx.doi.org/10.1016/S1098-2140\(99\)80206-9](http://dx.doi.org/10.1016/S1098-2140(99)80206-9)

Liem, A. (2018). Interview schedule development for a sequential explanatory mixed method design: Complementary-alternative medicine (CAM) study among Indonesian psychologists. *International Journal of Social Research Methodology*, 21(4), 513-525. <https://doi.org/10.1080/13645579.2018.1434864>

McGrail, M. R., Rickard, C. M., & Jones, R. (2006). Publish or perish: A systematic review of interventions to increase academic publication rates. *Higher Education Research & Development*, 25(1), 19-35. <https://doi.org/10.1080/07294360500453053>

Musal, B., Taskiran, C., Gursel, Y. Ü. C. E. L., Ozan, S., Timbil, S. E. V. G. İ., & Velipasaoglu, S. E. R. P. İ. L. (2008). An example of program evaluation project in undergraduate medical education. *Education for Health*, 21(1), 113. [https://www.researchgate.net/publication/23498646\\_An\\_example\\_of\\_](https://www.researchgate.net/publication/23498646_An_example_of_)



program\_evaluation\_project\_in\_undergraduate\_medical\_education

Nestel, D., Taylor, S., & Spender, Q. (2004). Evaluation of an inter-professional workshop to develop a psychosocial assessment and child-centred communication training programme for paediatricians in training. *BMC Medical Education*, 4(1), 1-10. <https://doi.org/10.1186/1472-6920-4-25>

Northcentral University. (2020). *Research process: Scholarly publication*. <https://ncu.libguides.com/researchprocess/scholarlypublication>

Owan, V. J., & Asuquo, M. E. (2022). "Publish or perish," "publish and perish": The Nigerian experience. In J. A. Undie, J. B. Babalola, B. A. Bello & I. N. Nwankwo (Eds.), *Management of higher education systems* (pp. 986-994). University of Calabar Press.

Owan, V. C., Basse, A. B., & Ubi, I. O. (2023). Construction and standardisation of an instrument measuring lecturers' persistence to publish in Scopus-indexed journals. *Journal of Applied Learning & Teaching*, 6(2), 158-171. <https://doi.org/10.37074/jalt.2023>

Scott, R. J., Cavana, R. Y., & Cameron, D. (2013). Evaluating immediate and long-term impacts of qualitative group model building workshops on participants' mental models. *System Dynamics Review*, 29(4), 216-236. <https://doi.org/10.1002/sdr.1505>

Shoko, A. P., Kimirei, I. A., Sekadende, B. C., Kische, M. A., & Sailale, I. E. (2021). Online course in conjunction with face-to-face workshops to improve writing skills leading towards more publications in peer reviewed journals. *European Science Editing*, 47, e54417. <https://doi.org/10.3897/ESE.2021.E54417>

Smidt, A., Balandin, S., Sigafos, J., & Reed, V. A. (2009). The Kirkpatrick model: A useful tool for evaluating training outcomes. *Journal of Intellectual and Developmental Disability*, 34(3), 266-274. <https://doi.org/10.1080/13668250903093125>

Steele, L. M., Mulhearn, T. J., Medeiros, K. E., Watts, L. L., Connelly, S., & Mumford, M. D. (2016). How do we know what works? A review and critique of current practices in ethics training evaluation. *Accountability in Research*, 23(6), 319-350. <http://dx.doi.org/10.1080/08989621.2016.1186547>

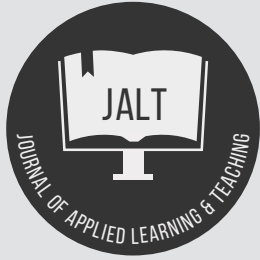
Sumathipala, A., Siribaddana, S., & Patel, V. (2004). Underrepresentation of developing countries in the research literature: Ethical issues arising from a survey of five leading medical journals. *BMC Medical Ethics*, 5(1), 1-6. <https://doi.org/10.1186/1472-6939-5-5>

Tella, A. (2015). *Publishing a good academic paper: My 16 years' experience*. A departmental seminar paper. Department of Library and Information Science, University of Ilorin.

Tella, A., & Onyancha, B. (2020). *Scholarly publishing experience of postgraduate students in Nigerian universities, accountability in research*. <https://doi.org/10.1080/08989621.2020.1843444>

Vinnikova, N., Kuzmenko, O., & Karpenko, A. (2021). Publication activity of Ukrainian higher academic staff: A case study. *Sabiedriba, Integracija, Izglitiba*, 6, 529-539. <https://doi.org/10.17770/sie2021vol6.6208>

Copyright: © 2024. Caroline Ochuko Alordiah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Professional development activities of English language lecturers in Vietnam through the lens of sociocultural theory

---

Le Pham Hoai Huong <sup>A</sup>	A	<i>Hue University, University of Foreign Languages and International Studies, Vietnam</i>
Tran Thi Thanh Thuong <sup>B</sup>	B	<i>Ho Chi Minh College of Foreign Economic Relations, Ho Chi Minh City, Vietnam</i>
Hung Phu Bui <sup>C</sup>	C	<i>University of Economics Ho Chi Minh City (UEH University), Vietnam</i>

---

### Keywords

---

English language lecturers;  
mediation;  
professional development (PD);  
sociocultural theory;  
teacher education.

---

### Correspondence

[lphuong@hueuni.edu.vn](mailto:lphuong@hueuni.edu.vn)<sup>A</sup>

---

### Article Info

Received 13 January 2024  
Received in revised form 25 February 2024  
Accepted 19 March 2024  
Available online 16 April 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.37>

### Abstract

---

Sociocultural theory emphasizes that social interaction and interaction with cultural artifacts lead to development. Adopting these principles of sociocultural theory, the current study aims to investigate the implementation of professional development activities for 56 English language lecturers at a tertiary institution in Vietnam and how these activities mediated their teaching careers. All of the lecturers were invited to participate in a survey, and 20 of them were interviewed. The findings indicate that for professional development, the English language lecturers considered interaction through students' feedback or course evaluation the most common activity, followed by interaction with colleagues via workshops, seminars, and conferences at different levels. In addition, the lecturers used videos, webinars, websites, books, and databases to create resources for their teaching. The mediating roles of social interaction, self-regulation, and artifacts were found to impact teachers' cognitive aspects, including knowledge, teaching skills, research, and also their enhanced motivation to work. From the findings, recommendations are put forward for the availability of regular training activities, opportunities for teachers to socialize, and the abundance of cultural and digital resources as affordances for teachers' professional development.

## Introduction

Professional development is part of any profession. In language teaching, professional development engages teachers in activities that lead to growth in their understanding and teaching skills. Professional development activities include those carried out at educational institutions or elsewhere, and they are generally grouped into activities such as learning, teaching, and researching done by individuals under the guidance of supervisors/mentors or in collaboration with others (Agbayahoun, 2016; Alshumaimeri & Almohaisen, 2017; Cirocki & Farrell, 2019). Examples of these activities include informal conversations with colleagues about teaching, English language teaching workshops, conferences, seminars, webinars, independent research, reading of professional materials, degree programmes, English language teaching training courses, and participating in class observations.

For professional development to take place, teachers are expected to participate in various activities to obtain knowledge and skills, which mediates their career growth. These activities must involve stakeholders, including school leaders and teachers, who must be informed of how to sustainably apply the knowledge and skills that they have obtained from the training courses (Agyei, 2022). Another recommended element for professional development includes developing teachers' bonding with colleagues within and outside their schools via social media platforms to build professional learning networks. The selection of learning networks can be based on their perceived weaknesses in content, pedagogy, or social-emotional attributes to improve their teaching skills (Mayeaux & Olivier, 2022).

This study adopts sociocultural theory proposed by Vygotsky (1978) and his colleagues as a theoretical framework to examine the mediation in the professional development of English language lecturers at a higher institution in Vietnam. According to sociocultural theory, human cognition is shaped and transformed through participating in social activities (Golombek & Doran, 2014). Mediation is the central concept of sociocultural theory, which refers to the developmental mechanism through which individuals, with time and interaction, via feedback and discussion with others, internalize knowledge and skills. Mediation in professional development can thus be realised and then internalized by individuals through observation and mentoring models given by others, such as teacher educators and more knowledgeable teachers. In other words, by engaging in professional development activities, teachers may go through a process whereby their understanding is refracted through the experience of others (Le, 2020), the use of cultural artifacts (e.g. learning resources) (Le & Bui, 2021), and self-regulation (Shi, 2017). Mediation can take place in most of daily and professional activities, but there remains a question of whether professional development activities can mediate language teachers in their institutional context. The overall focus of the current study is to explore the professional activities implemented by English language lecturers in Vietnam and to examine how mediation from doing these activities impacts their professional development. The findings of the current study thus can specify the mediation concept of sociocultural theory in teacher education and contribute to the application of the

concept in future studies in interpreting how growth in teaching professions can be realized. The findings can also provide practical implications for English language lecturers to develop themselves in terms of knowledge and skills in English teaching and for doing research required for the positions of university lecturers.

## Literature review and theoretical framework

### Sociocultural theory

Sociocultural theory emphasizes the social and cultural context in which mediation takes place via interaction in the learning process. Thanks to mediation, humans develop their thinking and are assisted in performing the activities they may not be able to do alone. One key concept in sociocultural theory is the zone of proximal development, which Vygotsky (1978) defines as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). The zone of proximal development has been interpreted to refer to assistance of different kinds, from interaction to the use of tools and artifacts leading to development in individuals. In teacher education, the zone of proximal development indicates the assistance teachers receive via talks with colleagues and experts or through interacting with digital tools to develop knowledge and skills related to their fields.

According to sociocultural theory, for development to take place, there must be "something that teachers do and enact in responding to pedagogical challenges and is mediated and influenced by a specific social context and conditions, and the interplay between these factors and individual efforts" (Ashton, 2021, p. 3) and "specific goal-oriented sociocultural activities" (Golombek & Doran, 2014, p. 104). In this study, these goal-oriented sociocultural activities are referred to as professional development activities in the teaching profession engaged by English language lecturers regarding learning, teaching, and researching. These activities can be carried out by themselves, under the guidance of more knowledgeable others (Le, 2007), or in collaboration with others.

### Sociocultural theory and professional development

This study adopted sociocultural theory as the theoretical framework to analyze the activities carried out by English language lecturers at an institution in Vietnam. More specifically, the concept of mediation in sociocultural theory guided the study. Mediation in sociocultural theory perspectives can be object-regulated, other-regulated, and self-regulated (Shi, 2017). First, to be object-regulated, EFL teachers are mediated by cognitive tools, such as English resources and social media, which serve to construct knowledge, consciousness, and reflection, which are considered to lead to professional development in teaching practice. Second, with regard to being other-regulated, assistance from others via communication, supervision, and interaction with other teachers, especially experts and

more capable teachers, is necessary to create new zones of professional development. Finally, to be self-regulated, English as a foreign language teachers respond to their work environment, reflect on their teaching, and manifest themselves to gain cognitive control for development in their profession. Sociocultural theory may also help explain professional development through a process of internalizing the external factors (Tasker et al., 2010). In other words, development in sociocultural theory is seen as the appropriateness of ideas that assimilate with pre-existing knowledge to reconstruct one's understanding. Accordingly, "internalization and transformation are individual, based on participation in social activities, and gauged by how these social activities are manifested" (p. 130). In quite a similar vein, Burner and Svendsen (2020) suggest considering teachers' backgrounds, and their use of tools and strategies to obtain changes and development through social interaction. Van Huizen et al. (2005) further specify the relationship between professional learning and development as the teachers' engagement in social events. They state:

Professional learning and development are best conceived and conditioned as an aspect of evolving participation in social practice. Participation involves being drawn into a setting that includes a program directed to the realization of values and goals, forms of social interaction and co-operation in an institutional context, and the use of cultural resources. (p. 274)

Seen from a sociocultural perspective, professional development takes place at certain times and in certain places with both constraints and affordances (White, 2018). In other words, viewing teacher professional development through the lens of sociocultural theory sheds light on how teachers develop their profession in their work settings as well as the activities that transform their understandings of the work of teachers and teaching practices (Tasker et al., 2010). Thus, sociocultural theory views teachers in the professional development process as taking an active role in creating their own knowledge and improvement (Khan et al., 2023). Besides, from the perspective of sociocultural theory, various means of mediation, such as tools and materials resources, signs, newsletters and journals, and other human activities, contribute to development in their profession.

### **Previous studies on using sociocultural theory as a framework to interpret professional development**

Studies on professional development from a sociocultural perspective have dwelled on the tenets of interaction, assistance, and how professional development is realized. For example, Kasi (2010) argued that the traditional model of knowledge transmission tends not to be beneficial to the professional development of EFL teachers. The author adopted sociocultural theory, specifically the concept of community of practice and collaborative action research, as an alternative model to actively involve novice teachers, experienced teachers, teacher-educators, and university researchers to pursue vibrant professional development in the field. The empirical studies on applying the new model

to the present situation of teaching and learning English were found to meet the need for change in the professional development of EFL in Pakistan.

Also inspired by the concept of community practice in sociocultural theory, Chen and Cheng (2014) explored teachers' professional growth in team teaching and the larger teaching context. Both Taiwanese and foreign English teachers paired the participants in their first year of teaching at the primary level in Taiwan. The findings of the study indicated favor for the situated and social nature of teacher learning. However, the socialization of the participants did not follow the typical learning process in a community of practice in which newcomers become more experienced members of a community.

Another professional development theme viewed through the lens of sociocultural theory is teacher agency development. In this direction, Kayi-Aydar (2019) investigated how a Hispanic language teacher invested in the agency while developing professional identities in different contexts. This case study adopted a narrative method with a recursive analysis of interviews and journal entries. The results showed three factors contributing to shaping the participants' agency and professional identity development - English language learning experience, discrimination and marginalization in the work environments, and knowledge gained through graduate studies. The author argued that the roles of ethnic and racial background, experience, and power differentials had links with the dynamic nature of teacher agency and identities.

In the same vein of exploring the concept of teacher agency in professional development, Uştuk and De Costa (2021) examined teachers' reflections on a lesson study model as a form of scaffolding. Field notes, interviews, and audio diaries were used to collect both introspective and retrospective data on reflective action both as a process and as a product. The findings indicated that reflective practice permeates lesson study as teachers reflect on their practice both individually and collaboratively. The study confirmed that teacher agency can be promoted in reflective practice throughout lesson study, and reflection might support transformative teacher agency among EFL teachers.

In a review of many professional development studies, especially those on teachers' online professional development between 2015 and 2019, Dille and Røkenes (2021) concluded that teachers' internal factors were crucial in their interactions with facilitators and peers. Among the factors, scaffolding was the most dominant category, with the main components as the teachers, their context, and online programs facilitating a shared understanding.

Uştuk and De Costa (2021) explored the nature of reflective practice in a professional development process among EFL teachers at a university in Turkey. Field notes, interviews, and audio diaries were used to collect both introspective and retrospective data. The findings showed that mediated by the lesson study model, the participants engaged in dialogic and collaborative teaching practice as a part of a professional development endeavor. They reinforced their teaching agency from the very beginning by choosing what

to work on and philosophy throughout the transformative process by developing new understanding and reframing their teaching practice.

Mediation also took the form of shared talks among teachers, as revealed in the studies by Zoshak (2016) and Shabani et al. (2010). For example, Zoshak (2016) explored “tiny talks” between teachers and colleagues about how they felt or what they struggled with could address an essential but often “extracurricular” aspect of teacher education. The findings revealed that “tiny talks” functioned as a mediational tool in transforming a change and facilitating progress in both being and becoming a teacher. Interactions in “tiny talks” fostered the casual nature of the interactions but implied possible applications to teacher education and professional development. Shabani et al. (2010) used Vygotsky’s zone of proximal development to explore instructional implications in teachers’ professional development. The study found that diary writing, peer and mentor collaboration, action research, practicum, and English language teaching discourse served as scaffolders to affect the progression of Zone of Proximal Development in language teachers. In another study, Eun (2008) found that professional development was a collaborative network that brought teachers together and a forum for teachers to share their intellectual challenges and resolutions, emotional struggles, and coping strategies. The study by Kuusisaari (2014) used the concept of the zone of proximal development to interpret the kinds of activities that support or hinder collaborative teacher learning during an in-service education course. The author pointed out that ideation, further development of ideas, and raising of questions led to collaborative development in teachers. However, excessive agreement appeared to prevent successful collaborative development.

Overall, the review of the related previous studies indicates that sociocultural theory can shed light on understanding mediation of English language teachers’ professional development. However, there are questions related to what specific forms of mediation take place when English language lecturers conduct professional development activities. Besides, it is worth focussing on one specific context, which is tertiary education in Vietnam, to explore what professional development activities are implemented by the English language lecturers at these institutions and how these activities impact or make changes to their careers. It is necessary to investigate these activities to provide insight into language teacher education and to promote professional development among the English language lecturers in Vietnam and in other contexts where the practice of professional development is similar. On these grounds, the current study was set out to answer for the following research questions:

- (1) What professional development activities are implemented by English language lecturers in Vietnam?
- (2) How do these activities mediate English language lecturers’ professional development?

## **Methodology**

### **Research design**

This study adopted a descriptive research design combining both qualitative and quantitative approaches. To be more specific, a questionnaire and an interview protocol were used to obtain data for the study. A 35-item questionnaire was designed, based on the perspectives of mediation from sociocultural theory. It consists of six clusters including (1) professional development activities implemented by English language lecturers via social interaction, (2) professional development activities implemented by English language lecturers via object-regulation, (3) professional development activities implemented by English language lecturers via self-regulation, (4) impact of professional development activities on English language lecturers via social interaction, (5) impact of professional development activities on English language lecturers via object-regulation, and (6) impact of professional development activities on English language lecturers via self-regulation were designed with the five-Likert scale collect data.

Besides, an interview protocol was designed with questions about the activities that lecturers had conducted themselves and with other colleagues/lecturers for professional development and their opinions about these activities, the resources they used, and the impact of these resources on their professional development.

### **Research setting**

The study was conducted at a university in Vietnam that trains English language teachers from primary to tertiary levels in the country. The lecturers in charge of teaching English at the institution received their master’s and Ph.D. degrees in teaching English to speakers of other languages or applied linguistics in Vietnam and other English-speaking countries. They have at least five years of experience in training English teachers in Vietnam. The participants in the current study were informed of the study’s purpose and consented to join the study.

### **Sample and data collection**

This study adopted convenient sampling to recruit voluntary participants. An invitation was sent to a tertiary institution in Vietnam via email and group contacts on social media, and a total of 56 English lecturers in Vietnam consented to take part in the study. For data collection, first, the participants completed a questionnaire and later, 20 volunteer lecturers of this population were chosen using the convenient sampling principle to take part in semi-structured interviews. All of the interviewees were anonymized in the data report and were referred to as T1-T20. It took about 10 minutes to complete a questionnaire and 20 minutes for each interview. The interview was semi-structured, with the questions prepared in advance, based on the mediation of sociocultural theory in relation to professional development. In the interviews, the researcher first explained the purpose of the study and then asked the participants to answer the questions in English

or Vietnamese. During the interviews, if further information was needed to clarify the lecturers' answers, then some probing questions such as "Please explain further" or "Please provide an explanation for your answer" were posed. On average, each interview lasted about 20 minutes, and the participants were put at ease to talk about their experiences implementing their professional development activities. The interview questions were to prompt the participants to recall their experiences and perceptions, and they were comfortable sharing since there were no controversial and ethical matters in the interviews.

## Data analysis

The qualitative and quantitative data were analyzed separately. The quantitative data collected from the questionnaire were first projected into the statistical package for the social sciences for statistical values, including the mean score and standard deviation of every single item before each factor's scale reliability (Cronbach's alpha) was examined. For data collected from the interviews, an inductive content-based process was employed to look for frequent, dominant, and significant themes emerging during the analysis (Mackey & Gass, 2005). The interview transcript coding was adapted from Tavil and Gungör (2017) and Shi (2017), given that the data from the interviews can highlight the themes related to sociocultural mediators and explain the growth through interactions with oneself, others, and objects. For example, the self-regulation themes had coded sub-themes as reading, attending conferences, and taking part in training courses; other-regulated themes included co-researching and co-teaching, and object-regulation themes included using artifacts for professional development.

## Findings

The questionnaire data was tabulated to see the result. The questionnaire's internal coefficient consistency (Cronbach's alpha) was 0.919, indicating its high reliability. First, the analysis of the questionnaire data for professional development activities implemented by English language lecturers via social interaction was processed (see Table 1).

In general, the activity achieving the highest mean score was the asking for students' feedback/course evaluation for professional development ( $M=3.8393$ ,  $SD=0.86921$ ), indicating that the participants' relatively high agreement with their reflections on their course evaluation at the end of their courses but their responses dispersed significantly on the five-level scale from the choice of rarely to always. This result could be explained by the fact that teachers were concerned about their students' opinions and course evaluations at the end of each semester to improve their teaching. It could also be explained that course evaluation is required at the research site. In comparison, the statement with the lowest mean score was "I collaborate with other English language lecturers to do research related to English language teaching for my professional development" ( $M = 3.3214$ ,  $SD = 0.97435$ ). Other statements about collaborating to design/prepare materials and lessons with colleagues

Table 1. Professional development activities implemented by English language lecturers via social interaction.

Statement	Min	Max	Mean	SD
1. I interact with other colleagues via workshops, colloquia, seminars, mentoring, etc. on topics related to English language teaching for my professional development.	2.00	5.00	3.7679	0.66033
2. I interact with other experts via workshops, colloquia, seminars, mentoring, etc. on topics related to English language teaching for my professional development.	2.00	5.00	3.5179	0.83101
3. I seek assistance from other English language lecturers through communication and cooperation with them for my professional development.	3.00	5.00	3.7321	0.70042
4. I seek supervision from experts or seniors capable English language lecturers for my professional development.	2.00	5.00	3.4107	0.75743
5. I join online discussion groups (Zalo, Facebook, forum, etc.) with other lecturers with the same interests for my professional development.	1.00	5.00	3.4643	1.09485
6. I collaborate to design/prepare materials and lessons with colleagues for my English teaching.	1.00	5.00	3.5000	0.93420
7. I ask for students' feedback/course evaluation for my professional development.	1.00	5.00	3.8393	0.86921
8. I collaborate with other English language lecturers to do research related to English language teaching for my professional development.	1.00	5.00	3.3214	0.97435

for English teaching had similar mean scores above 3.0, indicating the respondents' relatively high agreement with these professional development activities.

Besides, 19 out of 20 teachers interviewed stated that they took part in workshops, seminars, and conferences at different levels as professional development activities, indicating that the respondents' preferences for implementing these professional development activities or these activities were useful for them. For instance, T10 and T12 reported:

"I joined the short-term training courses to meet the immediate needs in my teaching job. Despite their short time, they are useful and help me to solve some problems in my teaching. I particularly appreciated the applicability and practicality of the course on how to teach phonics and pronunciation" (T 10).

"I could observe, listen to, and interact with the senior lecturers and other lecturers in workshops and seminars. I could also discuss this with other teachers to obtain more information related to teaching. More importantly, I could develop my teaching competence in the fields that I am concerned" (T 12).

Furthermore, co-researching with other colleagues was reported to be another activity carried out by the participants (eight teachers) and meeting other colleagues (six teachers). The implementation of these activities suggests that the teachers perceived the benefits of interacting and socializing in their professional domain. While teaching was mediated by interaction with mainly students and exchanging experiences with colleagues, researching was seen to be mediated by interaction with colleagues only. For example, one teacher said,

"I would like to share with colleagues more about the ideas and the research findings. We also talk about teaching methods, materials for teaching and researching, and how to deal with disturbing students in the classes. With young teachers who used to be my students, I also talk to them to learn from them new ideas in teaching and to give them advice" (T 1).

Teacher 1's response indicates that she understood the benefits of social interaction for her own teaching purposes and to support the younger ones in their professional activities.

In general, the professional development activities described by the participants were varied, profession-related, and carried out on a daily basis. They encompass interaction through different channels with students in the classroom, with colleagues in the institutional context, and with others in both online and offline, social and academic circles.

Regarding professional development activities via cultural tools, below is the finding for the professional development activities that the participants in the current study reported to be mediated by objects or artifacts.

Table 2. Professional development activities implemented by English language lecturers: Object-regulation.

Statement	Min	Max	Mean	SD
9. I use material resources/ databases provided by my university for my professional development.	2.00	5.00	3.9107	0.83724
10. I read professional materials, e.g. English language teaching magazines, journals, books or research findings in the field for my professional development.	3.00	5.00	3.8036	0.69856
11. I register for/pay for the applications (e.g. ELSA, Mindmeister, Kahoot, TED, etc.) for my professional development.	1.00	5.00	3.0000	1.06173
12. I use videos, webinars, websites, books and databases to obtain information for my English teaching.	2.00	5.00	4.2679	0.79752
13. I take teaching aids and other physical equipment as tools to transform external professional development knowledge into my own English teaching.	1.00	5.00	3.7500	0.89949

The use of the videos, webinars, websites, books and databases to have information to teach English achieved the highest agreement among the respondents (M=4.2679, SD=0.79752). On the contrary, the statement, "I register for/pay for the applications (e.g. ELSA, Mindmeister, Kahoot, and TED) for my professional development." received the lowest mean score (M=3.000), showing that the teachers tended to use the available resources and those material resources/databases provided by their university for their professional development (M =3.9107).

Besides, all the interviewed teachers said they used various resources and materials to mediate their teaching work. Those materials included both online and offline resources. For instance, T15 and T1 stated:

"I used Google Scholar and Google Book to find related materials and Sci-Hub to download documents." (T 15)

"The materials that I use include the online, man-made and artificial intelligence sources. The resources consist of the comments from my supervisors, colleagues' comments, and facial expressions. When I reflect on these input resources, I am able to internalize them to make them my own to serve my professional development". (T 1)

The training workshops were also considered to be the activities for professional development. Ten teachers confirmed the importance of joining the workshops as venues for obtaining knowledge and updating themselves with new teaching methods or information about using resources in teaching. For instance, one teacher said, "The training workshops provided me with knowledge and assessing tools to evaluate my students" (T 2). The other four teachers stated that the groups of similar professional concerns brought them ideas and resources for their career development.

Table 3. Professional development activities implemented by English language lecturers: Self-regulation.

Statement	Min	Max	Mean	SD
14. I reflect on my classroom teaching via narratives/teaching journals to strengthen my successes in teaching and to avoid the mistakes that I made before.	1.00	5.00	3.3036	1.00760
15. I self study topics related to my English teaching courses.	1.00	5.00	3.7679	0.97218
16. I visit other schools for professional development purposes.	1.00	5.00	2.6429	1.18212
17. I participate in training/research programs in other countries to enhance my professional development if given the opportunity.	1.00	5.00	3.1786	1.04633

Regarding the factors related to self-regulation for professional development, the quantitative data showed that the teachers expressed their high agreement with self-regulation (M=3.7679, SD=0.97218). However, their responses spread out in the five scales from strong disagreement to strong agreement. The statement "I visit other schools for professional development purposes." had the lowest mean value, at 2.6429. Other statements were in the range of 3.6 to 3.7.

In the interviews, most of the participants answered that they conducted self-regulation by reflection (10 teachers), doing action research (4 teachers), doing research (3 teachers), self-assessment (1 teacher), and reading research works related to their career specialization (2 teachers). Some teachers said:

"I observed my own teaching to know my strength." (T 2)

"Reflection is a cognitive activity. I do not write reflections but think about my teaching as a mental exercise and I am keen on doing such kind of reflections." (T3)

"I think about what I have taught and how to apply teaching methods more effectively. This has helped me improve my teaching everyday." (T 5)

The teachers' answers in the interviews indicate their high level of self-awareness of their professional development through the process of internalization using reflections because it is the procedure for them to adjust their teaching and improve themselves.

### Impact of mediated professional development activities on English language lecturers

Data about the impact of social interaction on their professional development collected from the survey with 56 English language lecturers is presented in Table 4 below:

Table 4. Impact of professional development activities on English language lecturers: Social interaction.

Statement	Min	Max	Mean	SD
18. Interacting with other colleagues via workshops, seminars, mentoring, etc. on topics related to English language teaching brings me knowledge and experience for my teaching career.	2.00	5.00	4.1786	0.54296
19. Interacting with other experts via workshops, seminars, mentoring, etc. on topics related to English language teaching makes me feel more confident in my teaching.	2.00	5.00	4.0179	0.70042
20. Seeking assistance from other English language lecturers is a way for me to be scaffolded to do things related to my career that I am not ready to do without their help.	2.00	5.00	4.0000	0.63246
21. Seeking supervision from experts or senior/more capable English language lecturers can support me to move to new levels of development in my career.	3.00	5.00	4.1250	0.63425
22. Joining online discussion groups (Zalo, Facebook, forums, etc.) with other lecturers of the same interests, I can broaden my knowledge related to my job.	2.00	5.00	3.7679	0.76256
23. Co-working with colleagues to prepare materials and lessons, we can mediate one another to design good lessons.	2.00	5.00	4.0714	0.62834
24. During the course and at the end of the course, I ask for students' feedback/course evaluation to adjust my teaching.	2.00	5.00	3.9286	0.80582
25. Collaborating with other English language lecturers to do research related to English language teaching brings me knowledge and skills in research that I cannot have if I work alone.	2.00	5.00	4.0357	0.76192

Generally, the statements in the table above achieved high mean values ranging from 3.7 to 4.17. The highest mean value was for the statement, "Interacting with other experts via workshops, seminars, mentoring, etc. on topics related to English language teaching makes me feel more confident in my teaching." which achieved a mean score of 4.1786 with a rather low standard deviation, indicating the high consistency of the teachers' responses (SD = 0.54296). However, the teachers agreed at the lower rate with the statement "Joining online discussion groups (Zalo, Facebook, forums, etc.) with other lecturers of the same interests, I can broaden my knowledge related to my job." (M = 3.7679). With regard to the impact of self-regulation on professional development, the data is presented in Table 5.

All the participants seemed aware of the importance of life-long learning when teaching English. The statement "I practise life-long learning to keep myself updated in English language teaching." reached the highest mean value of

Table 5. Impact of professional development activities on English language lecturers: Self-regulation.

Statement	Min	Max	Mean	SD
26. Reflecting on my classroom teaching actively improves my teaching skills.	2.00	5.00	3.9643	0.76192
27. By self-studying the topics related to my teaching courses, I am able to update my teaching knowledge and up-grading my teaching skills.	3.00	5.00	4.1964	0.51943
28. When writing teaching journals, I reflect on the strengths and weaknesses of teaching and have solutions to improve my teaching.	2.00	5.00	3.8036	0.77271
29. Visiting other schools for observing and teaching purposes is a way for me to learn from the others and to improve my teaching.	1.00	5.00	3.5357	0.83043
30. I purposefully join training/research programs in other countries to have opportunities to gain knowledge and skills that contribute to my professional development.	2.00	5.00	3.8929	0.70527
31. I practise life-long learning to keep myself updated in English language teaching.	3.00	5.00	4.3750	0.52440

4.3750 and a low SD of 0.52440, indicating the concentration toward the mean value in the teachers' responses. However, the statement "Visiting other schools for observing and teaching purposes is a way for me to learn from the others and to improve my teaching." generated a rather low mean score of 3.5357. Besides, the impact of "By self-studying the topics related to my teaching courses, I am able to update my teaching knowledge and up-grading my teaching skills." also achieved a high mean score of 4.1964.

Table 6. Impact of professional development activities on English language lecturers: Object-regulation.

Statement	Min	Max	Mean	SD
32. I gain knowledge for my teaching when using material resources/ databases provided by my university.	2.00	5.00	3.8571	0.67227
33. Reading professional materials, e.g. English language teaching magazines, journals, books, or research findings in the field makes me think more about the solutions for my teaching.	2.00	5.00	3.8929	0.80178
34. Thanks to digital tools, such as the internet, smartphones, and applications, I have made my teaching more effective.	3.00	5.00	4.4821	0.57179
35. The use of videos, webinars, websites, books and databases facilitates obtaining and transforming knowledge and skills for my professional development.	2.00	5.00	4.4464	0.71146

The teachers appreciated the usefulness of digital tools, for example, the internet, smart phones, and online applications to teach more effectively (M = 4.4821) as well as using videos, webinars, websites, books, and databases to facilitate the obtaining and transforming of knowledge and skills for my professional development (M = 4.4464). Other statements on the theme also obtained a relatively high mean value of 3.8, indicating the teachers' relatively high agreement with the use of tools and applications for professional development. Overall, the mean scores from 3.8 to 4.4 in Table 6 indicate the English language lecturers' understanding of the mediation regulated by the use of tools and artifacts for their professional development.

### Impact on teaching skills

In the interviews, half of the teachers described the impact of professional development on their professional knowledge. For example, one teacher said, "The reflections can help me to adjust my teaching, which is not logical, and realize students' needs." (T 4), and another stated, "Reflections bring me joy



and more ideas for teaching.” (T 18). One teacher answered in the interview, “I self-assessed my teaching strategies and the knowledge to be more confident in my teaching job as well as professional development.” (T 19). One teacher said that she wrote the teaching journal because “writing the teaching journal helped me to reflect more carefully and adjust my teaching activities. The internalization of the reflections makes me have critical thinking to teach more systematically.” (T 16). The participants’ answers indicate how they internalized the observation of their own teaching and made necessary changes to their teaching practice.

### **Impact on knowledge of teaching English**

The majority of the teachers answered in the interviews that participating in the training workshops or conferences brought them knowledge in their specialization, “I have obtained necessary knowledge and skills from the expertise of my colleagues and experts for my teaching career” (T 11). Another teacher stated,

“The professional development activities that I have done help me to reinforce my knowledge in certain fields, learn new ideas from colleagues, discuss some professional problems, and challenge my belief in implementing my professional development activities.” (T 13)

One teacher regarded interaction with other colleagues as one way to update professional knowledge, “Thanks to talking to other teachers, exchanging important information related to our profession, I have been able to improve my professional development” (T 19). Another teacher said that she gained more knowledge about her teaching from online materials: “I usually read materials on websites such as Edutopia, TeachThought, BookWidgets blog, Facebook community, v.v. These websites enlarge my understanding, share and receive teaching tips effectively.” Besides, professional knowledge is obtained not only from assistance and exchange with other colleagues but also from helping other teachers because “Assisting colleagues and students to gain knowledge and skills meaningfully is a way to improve my teaching skills” (T 20). The lecturers’ responses indicate the various venues they used for knowledge related to their teaching jobs. Among them, interaction with colleagues for professional development tended to be carried out more often than referring to online or offline resources.

### **Impact on research work**

In the interviews, eight teachers said that doing independent research or co-researching with other lecturers promoted their professionalism. One form of research that was mentioned by four interviewed teachers was action research. To be more specific, one lecturer said, “Action research helps me gain research competence to improve students’ learning in my classes.” (T 19). Other forms of research were also reported to mediate the lecturers. “What I have done with my colleagues promotes my research skills and enlarges knowledge in the field that I have not known before.” (T 03). Another lecturer also said, “I have collaborated with other

colleagues to develop research topics and teaching, which was useful for my lectures with graduate students. Research perspectives from Vietnam and other countries serve as useful input for my research and teaching work” (T 18). The participants’ answers reflect the requirements of doing research at the research site and that they were aware of this activity as part of professional development. They also indicated the importance of collaborating with colleagues for research work.

### **Impact on working motivation**

Implementing professional development activities, for example, designing lesson plans and materials was reported to give the teachers motivation and content in their job. “I feel contented with my teaching innovative ideas.” (T 15). The joy of teaching also came from sharing with other colleagues, “I received lots of advice and encouragement from other lecturers because they understand the difficulties that I encountered. Thanks to their encouragement, I was motivated to overcome the challenges and tried harder in my teaching career.” (T 14). According to another teacher, her motivation in work was from the internalization of the knowledge from other teachers and “other colleagues not only inspire me but also help me to find the right solutions for my teaching problems. Thanks to their support, I have become more creative and loved my career more.” (T 16).

In general, the responses from the English language lecturers indicate that they were mediated by social interaction with colleagues and other individuals from conferences and online forums and by interaction with tools and artifacts for knowledge and skills related to their teaching jobs. The impact of mediation was reported to benefit their teaching knowledge, skills, and motivation.

### **Discussion**

The current study aimed to document the professional development activities practiced by English lecturers at a tertiary institution in Vietnam and to explore how these activities impacted their professional development. It adopted sociocultural principles of mediation via social interaction and cultural artifacts to design the contents of the data collection tools, including a questionnaire and an interview protocol. Mediation in the current study explains how individuals internalize knowledge and skills from social interaction with more knowledgeable others or through the use of tools (e.g. digital tools, teaching resources) and develop their skills and understanding (Burner & Svendsen, 2020; Cirocki & Farrell, 2019; Eun, 2008). Seen from the sociocultural theory lens, the professional development activities reported in this study encompassed social interaction in the way that the teachers interacted with students via teaching and course feedback to improve their teaching and share professional knowledge with colleagues. It could be due to the institutional and contextual contexts of the current study, which required the teachers to provide regular teaching feedback that manifested the practice, leading to changes in the teaching practice and professional development. Besides, in the current study, the mediating

role of more experienced and skilled colleagues and experts was found to obtain high agreement among the participants, which confirms the importance of the critical concept of scaffolding in sociocultural theory (Sadeghi & Navaie, 2021; Vygotsky, 1978). Interaction with other colleagues and experts at workshops, conferences, and training courses was reported to mediate the professional development of the English lecturers in the current study.

Sociocultural theory also emphasizes the mediating role of cultural activities and artifacts (Le, 2020; Le & Bui, 2021). In this study, online materials and digital tools were found to mediate the thinking process, forming new ideas, knowledge, and skills related to the teachers' teaching. The teachers in the current study had the highest agreement on the use of videos, webinars, websites, books, and databases to gain knowledge for their teaching ( $M = 4.2679$ ). They also considered life-long learning important for professional development ( $M = 4.3750$ ). They indicated that they self-direct to gain an understanding of the topics related to their courses and enhanced teaching skills ( $M = 4.1964$ ) or did reflections on their professional development activities ( $M = 3.9643$ ).

Regarding the impacts of doing professional development activities, the current study found three emerging themes: impact on teaching skills and knowledge, which resonates with the findings of the studies by Agbayahoun (2016), Alshumaimeri and Almohaisen (2017), and Cirocki and Farrell (2019). Besides, social interaction was reported by the participants to help them gain knowledge related to their teaching, which reiterates the findings of the study by Li (2021). However, the current study documented data indicating the mediating roles of doing research in the affective domain, which has not been discussed. The teachers reported in the interviews that they obtained work motivation when interacting with colleagues. The teachers also stated that the assistance from online communities gave them motivation and sharing as a community of practice, as in the study by del Rosal et al. (2016). In other words, a more comprehensive understanding of the role of mediation via professional development activities was uncovered in the current study, shedding light on the roles of interaction for both cognitive and affective impact, using artifacts and participating in a community of practice related to the English teaching profession.

Central to the sociocultural theory is the concept of mediation which indicates the internalization of external factors including scaffolding from more knowledgeable others (Tasker et al., 2010) and the use of tools such as digital resources and databases (Le & Bui, 2021; Shi, 2017). This study specified this mediation process of sociocultural theory. It found that mediation was seen to lead to changes in teaching skills, knowledge of teaching English and doing research work for the professional development of the English language lecturers. For teaching skills, the teachers in the current study highly agreed with assistance from other English lecturers as a way to be scaffolded to do things related to their career that they were not ready to do without help. The participants also reported gaining knowledge from joining online discussion groups related to their teaching jobs and from interaction with other

lecturers of the same interests to broaden their knowledge of their careers. Moreover, using online materials and other resources assisted them with knowledge for teaching and researching, which is a requirement for university lecturers in Vietnam. In general, the mediation process tends to refer to the engagement of the English lecturers in various directions, including interaction with others, with materials, and with themselves.

## Conclusions and recommendations

In conclusion, from the findings of the current study on the mediating roles of social interaction, cultural activities, and artifacts in the professional development of English language lecturers, it is argued that the professional development of language teachers might necessarily be seen in connection with their colleagues, their students, and the affordances made available by the institutions and online communities. All these factors impact English language lecturers in terms of their teaching, research, and working motivation, to varying extent.

Sociocultural theory puts more emphasis on the mediating roles of social interaction and the use of tools, which leads to the development of thinking and teaching practice. This study's findings suggest that the mediating roles expand more than just the cognitive domain. Through interaction with colleagues, the teachers in the current study obtained their work motivation from sharing and encouragement, which is an indispensable factor in the teaching career. In other words, the current study contributes to the literature of sociocultural theory in terms of expanding the concept of mediation, which may lead to changes in teachers' affective factors related to their teaching careers. In this sense, professional development is more than just teaching, working, and researching. It is a more holistic picture of individual teachers in their work and their emotions.

From the findings of the current study, it is confirmed that the participant teachers appreciated professional development activities in any form and via any channel. Therefore, it is recommended that tertiary institutions provide teachers with opportunities to take part in these activities. Besides, interaction with colleagues in conferences or training workshops is considered to bring about enhanced knowledge and skills; thus, it might be essential that English teachers implement these professional development activities regularly. The emphasis on social interaction does not mean that autonomy, life-long, and self-study of teachers are not prioritized. Besides interaction with colleagues, English lecturers are recommended to learn more about and carry out self-mediated professional development activities, such as reflection. Reflections assist teachers to internalize teaching knowledge and skills. Furthermore, to realize the mediating roles of artifacts and digital tools, tertiary institutions might provide rich online and offline resources so that teachers can use them for professional development activities. In general, mediation via social interaction, self-mediation, or artifacts is meaningful and may contribute to transforming knowledge, understanding, and skills in teachers of English. Therefore, teachers and institutions can regularly implement professional development activities via

these platforms.

The current study was conducted at only one institution; therefore, the findings do not represent the professional development activities implemented in other educational contexts. Besides, the impacts of mediation via social interaction, self-regulation, and artifacts were explored by obtaining data from teachers' reflections and interviews. They were not measured statistically. Future studies can collect data quantitatively to document the quantified professional development of teachers.

## Acknowledgement

This work was supported by the research funding (Reference number: DHH2022-07-86) from Hue University, Vietnam.

## References

Agbayahoun, J. P. (2016). Teacher professional development: EFL teachers' experiences in the republic of Benin. *International Journal of English Linguistics*, 6(4), 144-152. <https://doi.org/10.5539/ijel.v6n4p144>

Agyei, D. D. (2022). Transferring and sustaining technological innovations after professional development: Insights from school leaders and teachers in Sub-Saharan Africa. *Journal of Applied Learning and Teaching*, 5(2), 31-40. <https://doi.org/10.37074/jalt.2022.5.S2.4>

Alshumaimeri, Y. A., & Almohaisen, F. M. (2017). Saudi EFL teachers' perceptions of professional development activities. *Journal of Education & Social Policy*, 7(1), 185-193. <https://doi.org/10.31235/osf.io/mw3e6>

Ashton, K. (2021). Novice teacher agency in the multi-level language classroom. *Language, Culture and Curriculum*, 34(3), 242-256. <https://doi.org/10.1080/07908318.2020.1818766>

Burner, T., & Svendsen, B. (2020). A Vygotskian perspective on teacher professional development. *Education*, 141(1), 11-20. [https://www.researchgate.net/publication/344677038\\_A\\_Vygotskian\\_perspective\\_on\\_teacher\\_professional\\_development](https://www.researchgate.net/publication/344677038_A_Vygotskian_perspective_on_teacher_professional_development)

Chen, C. W. Y., & Cheng, Y. S. (2014). Learning from team teaching and beyond: A case study on EFL teachers' professional development. *Journal of Pan-Pacific Association of Applied Linguistics*, 18(1), 33-47. <https://files.eric.ed.gov/fulltext/EJ1047439.pdf>

Cirocki, A., & Farrell, T. S. (2019). Professional development of secondary school EFL teachers: Voices from Indonesia. *System*, 85, 102111. <https://doi.org/10.1016/j.system.2019.102111>

del Rosal, K., Ware, P., & Montgomery, N. (2016). Mentoring teachers of English learners in an online community of practice. *International Journal of Computer-Assisted Language Learning and Teaching*, 6(3), 1-17. <http://doi.org/10.4018/IJCALLT.2016070101>

Dille, K. B., & Røkenes, F. M. (2021). Teachers' professional development in formal online communities: A scoping review. *Teaching and Teacher Education*, 105, 103431. <https://doi.org/10.1016/j.tate.2021.103431>

Eun, B. (2008). Making connections: Grounding professional development in the developmental theories of Vygotsky. *The Teacher Educator*, 43, 134-155. <http://doi.org/10.1080/08878730701838934>

Golombek, P., & Doran, M. (2014). Unifying cognition, emotion, and activity in language teacher professional development. *Teaching and Teacher Education*, 39, 102-111. <http://doi.org/10.1016/j.tate.2014.01.002>

Kasi, F. (2010). Collaborative action research: An alternative model for EFL teacher professional development in Pakistan. *Asian EFL Journal*, 12(3), 98-117. [https://www.researchgate.net/publication/260943717\\_Collaborative\\_Action\\_Research\\_An\\_Alternative\\_for\\_EFL\\_Teacher\\_Professional\\_Development\\_in\\_Pakistan](https://www.researchgate.net/publication/260943717_Collaborative_Action_Research_An_Alternative_for_EFL_Teacher_Professional_Development_in_Pakistan)

Kayi-Aydar, H. (2019). A language teacher's agency in the development of her professional identities: A narrative case study. *Journal of Latinos and Education*, 18(1), 4-18. <https://doi.org/10.1080/15348431.2017.1406360>

Khan, A. B., Ramanair, J., & Rethinasamy, S. (2023). Perceptions of Pakistani undergraduates and teachers of collaborative learning approaches in learning English. *Journal of Applied Learning and Teaching*, 6(1), 180-197. <https://doi.org/10.37074/jalt.2023.6.1.6>

Kuusisaari, H. (2014). Teachers at the zone of proximal development – collaboration promoting or hindering the development process. *Teaching and Teacher Education*, 43, 46-57. <https://doi.org/10.1016/j.tate.2014.06.001>

Le, P. H. H. (2007). The more knowledgeable peer, target language use, and group participation. *Canadian Modern Language Review*, 64(2), 329-350. <https://doi.org/10.3138/cmlr.64.2.329>

Le, P. H. H. (2020). The role of mediation in classroom interaction (Chapter 9, pp. 139-155). In H. Lee & B. Spolsky (Eds.), *Localizing global English: Asian perspectives and practices*. Routledge.

Le, P. H. H., & Bui, P. H. (2021). Mediation of digital tools in English learning. *LEARN Journal*, 14(2), 512-528. <https://so04.tci-thaijo.org/index.php/LEARN/index>

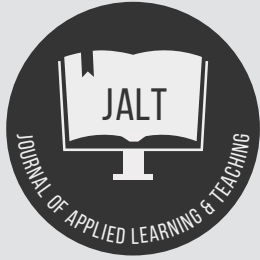
Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates Publishers.

Mayeaux, A. S., & Olivier, D. F. (2022). Professional kinship using social media tools: Bridging and bonding to develop teacher expertise. *Journal of Applied Learning and Teaching*, 5(1), 27-34. <https://doi.org/10.37074/jalt.2022.5.s1.3>

Sadeghi, K., & Navaie, L.A. (2021). Iranian EFL teachers' experiences with online professional development:

- Perceptions and preferences. *Iranian Journal of Language Teaching Research*, 9(3). <https://doi.org/10.30466/ijltr.2021.121073>
- Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English language Teaching*, 3(4), 237-248. [10.5539/elt.v3n4p237](https://doi.org/10.5539/elt.v3n4p237)
- Shi, H. (2017). The theoretical interpretation of EFL teacher's professional development from the perspective of sociocultural theory. *Theory and Practice in Language Studies*, 7(11), 1059-1064. <http://dx.doi.org/10.17507/tpls.0711.14>
- Tasker, T., Johnson, K. E., & Davis, T. S. (2010). A sociocultural analysis of teacher talk in inquiry-based professional development. *Language Teaching Research*, 14(2), 129-140. <https://doi.org/10.1177/1362168809353871>
- Tavil, Z. M., & Güngör, M. N. (2017). A sociocultural perspective on the development of Turkish pre-service teachers' competences and qualifications. *Pedagogy, Culture & Society*, 25(2), 263-277. <https://doi.org/10.1080/14681366.2016.1252788>
- Uştuk, Ö., & De Costa, P. I. (2021). Reflection as meta-action: Lesson study and EFL teacher professional development. *TESOL Journal*, 12(1), e00531. <https://doi.org/10.1002/tesj.531>
- van Huizen, P., van Oers, B., & Wubbels, T. (2005). A Vygotskian perspective on teacher education. *Journal of Curriculum Studies*, 37(3), 267-290. <http://dx.doi.org/10.1080/0022027042000328468>
- Vygotsky, L. S. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- White, C. (2018). Language teacher agency. In S. Mercer, & A. Kostoulas (Eds.), *Language teacher psychology* (pp. 196-209). Channel View Publications.
- Zoshak, R. (2016). 'Tiny talks' between colleagues: Brief narratives as mediation in teacher development. *Language Teaching Research*, 20(2), 209-222. <https://doi.org/10.1177/1362168815627659>

Copyright: © 2024. Le Pham Hoai Huong, Tran Thi Thanh Thuong and Hung Phu Bui. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Development and validation of an instrument to measure expectancy for success and subjective task value constructs in the context of higher education

---

Wan Hoong Wong<sup>A</sup>

A

*Singapore Institute of Management, Singapore*

Elaine Chapman<sup>B</sup>

B

*University of Western Australia, Perth, Australia*

---

### Keywords

---

Expectancy for success;  
Expectancy Value Theory;  
higher education;  
instrument development;  
subjective task value.

### Abstract

---

Given the relevance of the Expectancy Value Theory in the context of higher education, the aim of this study was to develop and validate an instrument to measure the constructs within this framework at the higher education level. Undergraduate students ( $n = 565$ ) from one of the largest private higher education institutions in Singapore were surveyed online using two versions (a 20-item and a 16-item version) of the *Expectancies and Values in Higher Education Instrument* (EVHEI). Exploratory factor analyses using a subsample of the cohort yielded two alternative versions of the instrument (a five-factor and a four-factor version). Both were subsequently validated using confirmatory factor analysis on data from the other subsample. The study results suggest that the EVHEI holds considerable promise for measuring motivation-related constructs at the higher education level.

---

### Correspondence

wongwanhoong@gmail.com <sup>A</sup>

---

### Article Info

Received 11 January 2024

Received in revised form 24 February 2024

Accepted 25 February 2024

Available online 5 March 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.20>

## Introduction

Motivation has long been regarded by many scholars and practitioners as a critical contributor to academic success at the college or higher education (HE) level (Lai, 2011). After reviewing both the psychological and the educational literature, Robbins et al. (2004) concluded that there is strong evidence that motivational constructs are linked consistently to college performance. The importance of motivation in educational outcomes has been emphasised to the extent that some researchers believe it would be untenable to draw inferences or make conclusions about educational outcomes without taking motivation into account. Similarly, Heckman (2007) stated that any evaluation of a human capital intervention would be seriously biased if motivation (alongside social adaptability) were excluded from the assessment due to an overemphasis on cognitive skills.

Numerous motivational theories, such as Bandura's (1997) self-efficacy theory (Doménech-Betoret et al., 2017), Deci and Ryan's self-determination theory (SDT) (Gagné & Deci, 2005), Weiner's attribution theory (Zhou & Urhahne, 2013), Covington's self-worth theory (van der Putten, 2017), and the expectancy-value theory of Eccles and her colleagues (Doménech-Betoret et al., 2017) have been proposed to operationalise the construct of motivation, particularly within education contexts. Educationists continue to draw from these popular theories to gain insights into how motivation relates to various other academic outcomes. Attribution theory, for instance, has been identified by Demetriou and Schmitz-Schiborski (2011) as the most widely applied motivation theory in the study of retention rates for undergraduate students.

In relation to HE, the role of motivation can be found in various theoretical or conceptual frameworks developed to explain students' academic achievement in their chosen courses of study. Motivation is included as an explicit component in the theoretical models of factors that influence academic performance by Allen (1999), Credé and Kuncel (2008), and Kusrkar et al. (2013). In the frameworks presented by Terenzini and Reason (2005) and Tinto and Pusser (2006), motivation is included as one of the student precollege characteristics that are associated most strongly with the persistence exhibited by HE students in their courses of study. The significance of motivation in explaining the academic success of HE students has also been tested in empirical research. For instance, motivation has been reported as a significant predictor of academic performance in HE students by Credé and Kuncel (2008), Griffin et al. (2012) and Morrow and Ackermann (2012).

## The Expectancy-Value Theory of achievement motivation

Research studies often use different motivational constructs underpinned by different motivation theories to examine the link between motivation and academic performance. One of the most important motivational frameworks within the literature on relationships between motivation and academic performance is the Expectancy-Value Theory (EVT) of Eccles and colleagues. This framework poses that

motivation beliefs relate to two key constructs, namely: (1) students' beliefs about how well they can perform certain tasks and (2) the values that the students attach to these tasks (Wigfield & Eccles, 2000). Steinmayr and Spinath (2009) argued that EVT was one of the "three most prominent" theories in the study of links between motivation and school achievement, a view that has been echoed in subsequent works (e.g., Gorges & Göke, 2015; Robbins et al., 2004).

On theoretical grounds, the EVT model is closely related to Bandura's (1997) concept of self-efficacy (Wigfield & Eccles, 2000; Wigfield et al., 2009). According to Wigfield and Eccles (2000), the expectancy construct in EVT is similar to the expectancy construct in Bandura's theory (which relates to task-specific expectancies), though not the outcome expectancy construct within the theory. It is important, however, to note that this similarity does not imply that self-efficacy and expectancies for success are indistinguishable. Wigfield et al. (2009) cautioned that while the constructs of ability beliefs, expectancies for success, and self-efficacy share similarities in their definitions, they are also distinct in important ways.

## Measuring the constructs within the expectancy-value framework

In the development of the EVT model, Eccles, Wigfield and their colleagues have included items to measure expectancies for success (EFS) and subjective task values (STVs), the two key constructs in the EVT framework (Wigfield & Eccles, 2000). The items developed by this research group have two specific features in common. First, all are related to a specific domain (mathematics). Second, they all target a specific group of learners (children). Perhaps due to these specificities, not all studies have used the original items developed by Eccles (1983). While their items have been used by Jones et al. (2010), other studies have measured expectancies for success and subjective task value either using items adapted from instruments previously developed by others (Bong, 2001; Chirinos, 2017; Dietrich et al., 2017; Doménech-Betoret et al., 2017; VanZile-Tamsen, 2001), or have created 'bespoke' measures specifically for use in their own studies (Battle & Wigfield, 2003; Gorges & Göke, 2015; Gorges & Kandler, 2012).

Table 1 provides a broad overview of how empirical studies within the literature have adapted and measured the constructs within the EVT framework. The table also demonstrates how the expectancies for success construct are measured typically in empirical studies. As indicated, numerous studies have used self-efficacy to operationalise the construct of expectancies of success in this body of work (e.g., Bong, 2001; Chirinos, 2017; Gorges & Göke, 2015; VanZile-Tamsen, 2001). Not all studies, however, treat self-efficacy and expectancies for success as similar constructs. For example, Doménech-Betoret et al. (2017) and Jones et al. (2005) included both constructs in their studies, measured separately with different sets of items.

Table 1. Measurement of the EVT constructs in past studies.

Study	EVT construct	Constructs developed for the purpose of the study	Construct measurement
Battle and Wigfield (2003)	Task value	Interest, attainment value, utility value and cost	Measured using items adapted from Eccles et al. (1983) and items created by the authors.
Bong (2001)	Expectancy for success	Self-efficacy for self-regulated learning, self-efficacy for academic achievement, content-specific self-efficacy and problem-specific self-efficacy	Measured using items taken directly or adapted from other studies (Zimmerman et al. (1992), Roser et al. (1996) and Pintrich and De Groot (1990)) and items created by the author.
	Task value	Perceived value of the course	Measured using three questions created by the author.
Chirinos (2017)	Expectancy for success	Mathematics self-efficacy	Measured using the 8-item scale from the 2012 PISA Student Questionnaire (mathematics self-efficacy).
	Task value	Task interest value and task utility value	Measured using 8-item scale from the 2012 PISA Student Questionnaire (intrinsic and instrumental motivation)
Dietrich et al. (2017)	Expectancy for success	Situational expectancies (expectations of success and competence beliefs)	Measured using two items adapted from Wigfield and Eccles (2000) and Schneider et al. (2015)
	Task value	Situational task values (intrinsic value, utility value for future job, personal importance, effort cost, emotional cost and opportunity cost)	Measured using seven items adapted from Gaspard et al. (2015)
Domènech-Betoret et al. (2017)	Expectancy for success	Cost expectancy, achievement expectancy and process expectancy	Measured using 13 items adapted from Domènech (2006, 2012, 2013) and Domènech-Betoret et al. (2014) (ten items for the three constructs of expectancy of success and three items for the task value constructs)
	Task value	Subject value	
Study	EVT's construct	Constructs developed for the purpose of the study	Construct measurement
Jones et al. (2010)	Expectancy for success	Expectancy for success in engineering	Measured using 2-item scale adapted from Eccles and Wigfield (1995)
	Task value	Engineering intrinsic value, engineering attainment value and engineering extrinsic value	Measured using 3 scales with total of 7 items adapted from Eccles and Wigfield (1995)
Gorges and Goke (2015)	Expectancy for success	Task-specific (novel academic tasks) self-efficacy	Measured using three items created by the authors
Gorges and Kandler (2012)	Expectancy for success	Current expectation of success	Measured using nine items created by the authors
	Task value	Current value	Measured using nine items created by the authors
VanZile-Tamsen (2001)	Expectancy for success	Expectancy of success	Measured using items from Motivated Strategies for Learning Questionnaires (MSLQ) which relate to self-efficacy and attributions of success and failure.
	Task value	Task value	Measured using items from Motivated Strategies for Learning Questionnaires (MSLQ) which relate to intrinsic goal orientation and task value.

## Expectancy-Value Theory and academic performance in higher education

As noted, the EVT model focuses on the ability beliefs and subjective task values of children and adolescents (Wigfield & Eccles, 2000). A number of studies have reported significant relationships between the components of the EVT model (i.e. expectancies for success and subjective task values) and academic performance in this population. For example, Wigfield and Eccles (2000) reported ability beliefs and expectancies for success in children to be one of the strongest predictors of grades in mathematics. Chirinos (2017) reported that efficacy beliefs among Latino high school students predicted their academic behaviours and performance in mathematics. Similarly, Steinmayr and Spinath (2009) found that ability self-perceptions and values (alongside other motivational constructs) contributed to the prediction of school achievement in 11th- and 12th-grade students. Reviews by Chirinos (2017), Wigfield and Eccles (2000) and Wigfield et al. (2009) also highlighted numerous other empirical studies that demonstrated the relevance of the EVT model in explaining the academic performance outcomes of children and adolescents.

Research applying the EVT model in studies with adult learners, such as students in HE, are much fewer than studies with younger learners such as children or adolescents (Gorges, 2015; Gorges & Kandler, 2012). There is, however, some empirical evidence to support the claim that the EVT model can also be relevant to adult learners. For instance, through a meta-analysis of 109 studies, Robbins et al. (2004) concluded that academic self-efficacy (which was linked to the EVT construct of expectancies) was the best predictor of GPAs in college students. Chirinos (2017) and Wigfield et al. (2009) also reviewed and cited a few studies that applied the EVT framework in exploring the academic performance of HE students. Bong's (2001) and Chirinos' (2017) studies, for example, found that components of the EVT model were able to predict, to a moderate extent, academic performance in HE settings.

The relevance of the EVT model in the context of adult learners is not limited to academic performance. Expectancies for success and/or subjective task values are also found to be significantly related to other variables in HE such as enrolment intentions (Bong, 2001), career aspirations (Battle & Wigfield, 2003; Jones et al., 2010), use of self-regulated strategies (VanZile-Tamsen, 2001), levels of motivation to use new learning opportunities (Gorges & Kandler, 2012) and the degree of effort that students expend on their studies (Dietrich et al., 2017).

In addition, it is important to consider the different dimensions of subjective task values in discussing the relevance of EVT in the context of HE. Unlike expectancies for success, which are conceptualised as a unidimensional construct, subjective task values comprise four different dimensions – attainment value, utility value, intrinsic value and cost (Wigfield & Eccles, 2000). From the literature, however, it appears that the role of cost as a constituent component of STV is ambiguous. While cost is associated with the effort to accomplish an activity in the EVT framework (Wigfield & Eccles, 2000), Gorges (2015) argued that cost should be a separate construct.

In the same vein, Barron and Hulleman (2015) proposed that cost should be treated as a distinct motivational construct from expectancies and values, and thus established their Expectancy-Value-Cost model. This is unsurprising, given the lack of empirical support for the cost construct as theorised in the original EVT framework. In discussing the framework, Wigfield and Eccles (2000) pointed out that most of their empirical work in relation to the framework had been mainly focused on the other three task values and not cost.

Evidence supporting the notion that cost may be a separate motivational construct altogether was reported in a study of 155 German university students (Dietrich et al., 2017). In this study, the relationship between expectancies, task values and student effort was examined. The authors reported that the fit of the measurement model was superior with cost treated as a separate construct from the expectancies and subjective task values constructs.

## Rationale and aims of the present research

While the EVT model has thus far been applied predominantly at the primary and secondary levels, previous research has confirmed its potential utility in predicting academic behaviours and outcomes at the HE level. Despite this, a generic instrument to measure the expectancy-value constructs in the context of HE success is not yet available. While instruments measuring the EVT constructs have been published, most of these are intended for use with younger learners. These instruments will not be applicable in the context of HE, because in students at this higher level, the elements of both task values and cost would need to be operationalised in a very different way, though that of expectancies may be more similar (Sogunro, 2015; Yoo & Huang, 2013).

Instruments suitable for measuring the EVT constructs at the HE level are limited. Furthermore, items in the existing instruments that have been developed for use at this level

have referenced outcomes such as learning effort (Dietrich et al., 2017), novel academic tasks (Gorges & Göke, 2015), career plans (Jones et al., 2010) and students' self-regulated strategy use (VanZile-Tamsen, 2001), rather than academic success. These instruments have also measured expectancies for success as self-efficacy, despite the fact that the two constructs are conceptually distinct (Wigfield et al., 2009).

Further to the above points, the instruments developed thus far to measure the EVT constructs in the context of HE have typically focused on selected elements of the model. Two studies, for instance, focused only on developing and validating items related to the STV dimensions of the EVT model. In a study conducted on a group of post-undergraduate students from six institutions in the United States (Brunhaver et al., 2017), a 15-item instrument to measure the STV components of the EVT model was developed. The exploratory factor analyses (EFAs) and confirmatory factor analyses (CFAs) indicated a three-factor solution, which corresponded to the elements of attainment value, intrinsic value and cost. In another study conducted on a group of students from a public university in the United States (Flake et al., 2015), a 19-item instrument to measure the cost component of the STV construct was developed. The EFA and CFA conducted indicated a four-factor solution, which corresponded to the EVT elements of task effort cost, outside effort cost, costs associated with the loss of valued alternatives, and emotional cost.

In view of the relevance of the EVT constructs in the context of HE, the aim of the present study was to develop a stand-alone instrument to measure the EFS and different dimensions of the STV constructs within the EVT model, with specific reference to academic success in HE. Unlike most existing instruments, the instrument in this study related to academic success in a broader sense, rather than within a specific domain. This was done to ensure the general utility of the instrument across academic HE contexts. The authors were of the view that, in the context of HE, the key concern of learners is not generally about their performance in a specific domain area, but about their performance in a more general sense. As a result, the items in the instrument developed all referred to the respondents' ability to complete and graduate from their chosen HE programmes.

## Method

### Participants and settings

Participants were students from one of the largest private HE institutions in Singapore, which offered 14 international undergraduate degree programmes taught by an institution from the United Kingdom. For the purpose of this study, the students were invited to participate in an online survey. In total, 565 of the students responded and completed the survey. Of this sample, 219 (38.76%) of the participants were males and 346 (61.24%) were females. The ages of the respondents ranged from 17 to 30 years (mean age 22.05 years,  $SD=2.08$ ). In terms of nationality, 397 (70.27%) were students from Singapore, 107 (18.94%) were from other Southeast Asian countries, 57 (10.09%) were from other Asian countries, and 4 (0.71%) were from countries outside

of Asia.

### Instrument development

Based on the EVT model developed by Eccles, Wigfield and their colleagues (Wigfield & Eccles, 2000), the Expectancies and Values in Higher Education Instrument (EVHEI) developed in this study comprised two scales – the Expectancy for Success (EFS) scale and Subjective Task Value (STV) scale. These two scales were intended to measure the two key constructs of *expectancies for success* and *subjective task values*, in the context of achieving academic success in HE. The EFS component included no subscales and comprised four items. The STV component, on the other hand, was designed to incorporate four subscales: *Attainment Value (AV)*, *Utility Value (UV)*, *Intrinsic Value (IV)* and *Cost (CST)*. Each subscale comprised four items. Thus, 20 items were created, each of which was presented in the form of a 7-point bipolar statement rating scale. Respondents were required to select a point on the scale which best described their own position with respect to the two polar statements. For each item, scores ranged from 1 to 7. The item statements corresponding to the EFS and STV scales in the EVHEI are shown in Table 2.

Given the ambiguous role of cost as an integral component of the STV, two alternative versions of the EVHEI were developed and validated. Version 1 comprised all 20 items measuring the EFS, AV, UV, IV and CTS subscales; while Version 2 comprised only 16 items, measuring the EFS, AV, UV and IV subscales (i.e., with the four cost or CST items excluded).

Table 2. Item statements in the Expectancy-Value for Higher Education Instrument (EVHEI).

Scale	Subscale	Item Label	Bipolar Item Statement (Lowest score = 1 vs. Highest score = 7)		
Expectancy for Success	Nil	EFS1	Score = 1 I feel it is very unlikely that I will be able to do well in the courses I am currently taking. Score = 7 I have complete confidence that I can do well in the courses I am currently taking.		
		EFS2	Score = 1 I am very doubtful that I can progress to the next academic year. Score = 7 I have no doubts at all that I can progress to the next academic year.		
		EFS3	Score = 1 I am very doubtful that I can pass all the courses I am currently taking. Score = 7 I have no doubts at all that I can pass all the courses I am currently taking.		
		EFS4	Score = 1 I am very uncertain of whether I can graduate from this programme. Score = 7 I am very certain that I can graduate from this programme.		
	Subjective Task Value	Attainment Value (AV)	AV1	Score = 1 It is not at all important for me to do well in my studies. Score = 7 It is extremely important for me to do well in my studies.	
			AV2	Score = 1 It is not at all important for me to obtain university-level education. Score = 7 It is extremely important for me to obtain university-level education.	
			AV3	Score = 1 It is not at all important for me to learn the knowledge and skills from the courses I am taking. Score = 7 It is extremely important for me to learn the knowledge and skills from the courses I am taking.	
			AV4	Score = 1 It is not at all important for me to graduate from this BSc undergraduate programme. Score = 7 It is extremely important for me to graduate from this BSc undergraduate degree programme.	
Utility Value (UV)		UV1	Score = 1 Getting a higher salary is not at all the reason why I undertook the programme. Score = 7 Getting a higher salary is an extremely important reason why I have undertaken this programme.		
		UV2	Score = 1 Securing a better future career prospect is not at all a reason why I undertook the programme. Score = 7 Securing a better future career prospect is an extremely important reason why I undertook the programme.		
		UV3	Score = 1 Having a more successful life is not at all an extremely important reason why I undertook the programme. Score = 7 Having a more successful life is not at all a reason why I undertook the programme. [Note: Success can be material or non-material oriented]		
		UV4	Score = 1 Getting recognition from society is not at all a reason why I undertook the programme. Score = 7 Getting recognition from society is an extremely important reason why I undertook the programme.		
		Subjective Task Value	Intrinsic Value (IV)	IV1	Score = 1 Undertaking this programme has nothing to do with my own enjoyment. Score = 7 I am undertaking this programme largely for my own enjoyment.
				IV2	Score = 1 Getting personal satisfaction is not at all a reason why I undertook the programme. Score = 7 Getting personal satisfaction is an extremely important reason why I undertook the programme.
				IV3	Score = 1 Getting a sense of achievement is not at all a reason why I undertook the programme. Score = 7 Getting a sense of achievement is an extremely important reason why I undertook the programme.
				IV4	Score = 1 Becoming more knowledgeable is not at all a reason why I undertook the programme. Score = 7 Becoming more knowledgeable is an extremely important reason why I undertook the programme.
Cost (CST)	CST1		Score = 1 There is nothing to give up at all for me to undertake this programme. Score = 7 I give up a great deal of my life by undertaking this programme.		
	CST2		Score = 1 I absolutely have no other better things to do if I don't attend classes. Score = 7 I definitely have other better things to do if I don't attend classes.		
	CST3		Score = 1 There is no change at all in my stress level since I took up this programme. Score = 7 My stress level has increased tremendously since I took up this programme.		
	CST4		Score = 1 My mental energy remains the same after attending a 3-hour lecture. Score = 7 My mental energy has drained completely after attending a 3-hour lecture.		

### Procedure

The questionnaire survey was administered online through the Qualtrics platform. Students were invited via e-mail to participate in the survey on a voluntary basis at the beginning



of the 2018-2019 academic year. The purpose of the survey, the time required to answer the survey, the confidential nature of the survey, and data protection assurances were also included in the e-mail. Participants were required to consent using a radio button before they proceeded with the online survey. Following the initial invitation, two e-mail reminders were sent to increase the participation rate.

Prior to the actual survey, a pilot study was conducted with a small group of students ( $n = 14$ ). The purpose of this step was to assess the clarity of the instructions, the suitability and clarity of the questions, and the time required to complete the online version of the survey. Written feedback was obtained from the participants. Results showed that the instructions and questions were clear and appropriate and that the indicated time of 15 minutes to complete the survey was reasonable. Given that no major issues were identified in this pilot study, only minor amendments were made to the questionnaire to improve it before the final launch.

### Data analysis using exploratory factor analysis and confirmatory factor analysis

Both EFAs and CFAs were conducted to evaluate and validate the internal structure of the two EVHEI versions. The sample was first randomly split (using a random number generator) into two approximately equal-sized subsamples. Based on this random split approach, Subsample 1 included 265 observations and Subsample 2 included 300 observations. The factor analyses were then conducted in two stages, with EFA first performed using Subsample 1 followed by CFAs using Subsample 2. In each stage, the subsample was used to evaluate both versions of the instrument (the full 20-item version and the 16-item version, which excluded the cost dimension).

The rationale for using EFAs in conjunction with CFAs in testing newly created items has been provided by Brown (2006), Fabrigar and Wegener (2011), Osborne (2014), Post and Walma van Der Molen (2019) and Yong and Pearce (2011). Although the EVHEI was intended to measure the two key constructs, as defined by the established EVT model, its items were newly created. In addition, the context involved (i.e. academic success in HE) was new. Therefore, EFAs were first conducted to provide a preliminary assessment of the underlying factor structure of the newly created items. This was then followed by CFAs to confirm the factor structures derived from the EFAs. Such an approach is commonly adopted in instrument validation studies (Jansen et al., 2017; Post & Walma van Der Molen, 2019). For the purpose of this study, the EFAs were conducted using SPSS V24, while the CFAs were conducted using LISREL V10.20.

## Results

Preliminary data screening analyses indicated several missing responses and the presence of outliers, which were subsequently removed from the dataset. This resulted in 246 cases being retained for Subsample 1 and 277 cases being retained for Subsample 2. Checks for normality, linearity, the presence of multicollinearity and factorability were also

conducted on the two datasets. No apparent violations of these requirements were found (see Table 3). Descriptive statistics for all of the EVHEI's 20 items by subsample are provided in Table 4.

Table 3. Preliminary examination of the data prior to the factor analyses.

	Subsample 1 for EFA ( $n = 265$ )	Subsample 2 for CFA ( $n = 300$ )
Removal of missing responses and multivariate outliers identified using Mahalanobis distance	Four cases with missing responses (1.5%), with no systematic pattern. These cases were removed (see Tabachnick & Fidell, 2013). 15 cases (5.7%) considered outliers with $\chi^2$ statistic less than the critical value at 0.001 significance level. These cases were removed. A total 246 cases were retained after the removal with case to ratio 15:1 (recommended 10:1; see Yong and Pearce (2013)).	Three cases with missing responses (1%), with no systematic pattern. These cases were removed. 20 cases (6.7%) considered outliers with $\chi^2$ statistic less than the critical value at 0.001 significance level. These cases were removed. A total of 277 cases were retained after the removal with case to ratio 17:1.
Assumption testing		
1. Normality	All the items demonstrated some level of skew based on skewness and kurtosis coefficients, Kolmogorov-Smirnov and Shapiro-Wilk tests. This violation should not compromise the outcomes of the EFA and CFA analyses.	
2. Linearity	Visual examination of bivariate scatter plot revealed no non-linear relationships.	
3. Multicollinearity	Examination of correlation matrix revealed all the correlation coefficients were well below 0.90, a recommended guideline (see Tabachnick & Fidell, 2013; Yong & Pearce, 2011). This indicates no serious problem of multicollinearity.	
4. Factorability (EFA only)	The Kaiser-Meyer-Olkin sampling adequacy measure of 0.88 is well above the recommended value of above 0.50 (see Yong & Pearce, 2013) or above 0.60 (see Tabachnick & Fidell, 2013). Bartlett's test was significant with $p$ -value < 0.001. This indicates that EFA was a tenable analysis approach to be used with this dataset.	

Table 4. Descriptive statistics for Subsample 1 (EFA) and Subsample 2 (CFA).

Item	Subsample 1 ( $n = 246$ )	Subsample 2 ( $n = 277$ )
	$M(SD)$	$M(SD)$
EFS1	4.87 (1.22)	4.87 (1.27)
EFS2	5.37 (1.34)	5.31 (1.49)
EFS3	5.22 (1.40)	5.25 (1.48)
EFS4	5.63 (1.43)	5.55 (1.63)
AV1	6.23 (1.13)	6.22 (1.13)
AV2	6.28 (1.18)	6.33 (1.11)
AV3	6.24 (1.10)	6.11 (1.13)
AV4	6.37 (1.14)	6.38 (1.06)
UV1	5.31 (1.35)	5.24 (1.48)
UV2	5.95 (1.23)	5.93 (1.25)
UV3	5.75 (1.35)	5.77 (1.28)
UV4	4.97 (1.59)	4.92 (1.67)
IV1	4.59 (1.46)	4.65 (1.60)
IV2	4.88 (1.44)	5.04 (1.47)
IV3	5.23 (1.42)	5.31 (1.45)
IV4	5.75 (1.24)	5.89 (1.09)
CST1	4.19 (1.45)	4.36 (1.67)
CST2	3.93 (1.57)	3.89 (1.68)
CST3	4.83 (1.41)	5.00 (1.49)
CST4	4.75 (1.45)	4.83 (1.51)

### Exploratory factor analysis

Given that the assumption of multivariate normality was violated, Principal Axis Factoring (PAF) was chosen as the method of extraction instead of Maximum Likelihood (ML). The principal factor method is regarded to be a more suitable method of extraction when the assumption of multivariate normality has not been met (Fabrigar et al., 1999; Osborne, 2014; Yong & Pearce, 2013). The factors were then rotated to approximate simple structure using an oblique rotation method (Direct Oblimin), given the likelihood that these would be correlated (Osborne, 2014; Tabachnick & Fidell, 2013; Yong & Pearce, 2013). For example, in the EVT model,

the constructs of expectancies for success and subjective task values are both deemed to be influenced by goals and self-schemata (Wigfield & Eccles, 2000).

The initial EFAs were conducted both with all 20 items of the instrument included (Version 1) and with the cost items removed (i.e., 16 items - CST1 to CST4 excluded – Version 2). Using Kaiser’s criterion, the EFAs indicated five distinct factors for Version 1 and four distinct factors for Version 2. Table 5 provides an overview of the EFA conducted for the two versions.

Table 5. Comparative EFAs for the two instrument versions (n = 246).

Dimensions considered	Version 1	Version 2
	EFS, AV, UV, IV	EFS, AV, UV and IV and CST
Number of items	20	16
Number of factors extracted	5	4
Proportion of the item variation accounted for by the factors extracted	73.98%	80.11%
Range of extracted communalities	.14 – .92	.52 – .92

Table 6 presents the factor loadings for Versions 1 and 2 of the EVHEI. In determining the factor structure of an instrument, Matsunaga (2010) regarded .40 as the lowest acceptable loading. This cutoff was adopted by Battle and Wigfield (2003) in assessing the factor structure of the EVT task value construct in college women’s value orientations.

Based on a cutoff threshold of .40, the pattern matrices in Table 6 show a clear factor structure for both versions of the instrument. In both, no item was cross-loaded notably onto two or more factors, and items clustered together as expected based on the EVT constructs. In Version 1, the 20 items loaded on five factors with EFS1 to EFS4; AV1 to AV4; UV1 to UV4; IV1 to IV4; and CST1 to CST4 loaded unambiguously onto five separate factors. The loading of CST2 was, however, somewhat lower than .40, suggesting that this particular item was more weakly associated with others in the cost factor. In Version 2, the 16 items loaded on four factors with EFS1 to EFS4; AV1 to AV4; UV1 to UV4; and IV1 to IV4 loaded unambiguously on four separate factors.

Overall, the factor structures obtained were consistent with the theoretical framework of the EVT. As expected, the results also indicated that the factors were moderately correlated, with correlation coefficients in the range of .01 to .50 for Version 1, and .21 to .53 for Version 2.

Table 6. Factor loading of items in the EVHEI for Versions 1 and 2 (n = 246).

Item	Communalities (h <sup>2</sup> )	Version 1 (20 items)					Version 2 (16 items)												
		Pattern Matrix		Structure Matrix			Pattern Matrix		Structure Matrix										
		Factor 1	2	3	4	5	Factor 1	2	3	4	5								
EFS1	.68	-.73	.02	.42	.17	-.05	-.80	.22	.21	.45	-.06	-.76	-.04	.02	.13	.81	.31	.20	.45
EFS2	.92	-.92	.13	.25	.02	-.01	-.95	.30	.23	.41	-.03	.95	.05	.00	-.03	.96	.39	.22	.40
EFS3	.90	-.93	.05	.20	.01	-.04	-.95	.21	.18	.37	-.06	.99	-.05	-.01	-.04	.95	.31	.17	.37
EFS4	.75	-.82	.11	.02	.05	.04	-.86	.27	.22	.40	.03	.85	.02	.01	.01	.87	.36	.21	.40
AV1	.72	-.14	.78	.05	.03	-.02	-.28	.83	.47	.38	-.04	.05	.83	-.00	.00	.37	.85	.45	.36
AV2	.80	-.07	.81	.13	.00	-.03	-.23	.89	.54	.37	-.04	-.02	.87	.07	-.02	.32	.90	.53	.35
AV3	.74	-.09	.73	.05	.17	.01	-.28	.83	.50	.48	-.01	.01	.79	.02	.14	.37	.85	.48	.46
AV4	.88	-.07	.91	.01	.01	.08	-.23	.93	.50	.38	.05	-.02	.95	-.02	-.01	.33	.93	.48	.35
UV1	.61	-.07	-.04	.82	-.11	.01	-.15	.34	.77	.20	.13	.08	-.04	.81	-.09	.20	.38	.78	.19
UV2	.75	-.07	.81	.13	.00	-.03	-.23	.89	.54	.37	-.04	.05	.21	.75	-.09	.25	.60	.85	.26
UV3	.73	-.04	.16	.75	.01	-.02	-.18	.55	.84	.36	.09	.03	.18	.73	.01	.26	.39	.83	.34
UV4	.53	.10	-.11	.69	.22	.01	-.07	.30	.70	.39	.12	-.09	-.09	.69	.22	.12	.34	.70	.38
IV1	.52	-.06	.01	-.09	.72	.05	-.31	.23	.19	.71	.02	.06	.01	-.07	.71	.36	.28	.18	.72
IV2	.80	-.03	-.05	-.01	.96	-.02	-.37	.30	.32	.95	-.01	.03	-.03	.00	.95	.43	.36	.31	.95
IV3	.71	-.02	.05	.14	.75	.01	-.32	.40	.44	.83	.04	.01	.38	.14	.73	.39	.46	.43	.82
IV4	.66	-.11	.28	.10	.55	.02	-.39	.55	.46	.74	.03	.11	.30	.09	.53	.47	.60	.45	.72
CST1	.21	-.12	-.13	.06	.07	.41	-.12	-.07	.10	.09	.42	-	-	-	-	-	-	-	-
CST2	.14	-.10	-.14	-.06	-.09	.03	-.03	-.20	-.11	.12	.29	-	-	-	-	-	-	-	-
CST3	.54	.25	.23	.08	.03	.64	.21	.21	.25	.05	.65	-	-	-	-	-	-	-	-
CST4	.41	.11	.14	.01	.02	.61	.09	.11	.16	.04	.61	-	-	-	-	-	-	-	-

The internal consistency of the items was evaluated using Cronbach’s α coefficients, as shown in Table 7. Ho (2014) suggested that a high internal consistency is attained when the Cronbach’s α coefficient is greater than .80. While the Cronbach’s α coefficients for EFS, AV, UV and IV scales were all above .80, the Cronbach’s α coefficient for CST (.54) was noticeably lower than this threshold. This suggests that the internal consistency for the CST scale was weak, particularly in comparison to the other subscales. In light of this, the overall Cronbach’s α coefficient for the 20-item instrument was lower with the four CST items included (.87) than for the overall Cronbach’s α coefficient with the four CST items removed (.91).

Table 7. Cronbach’s α of the EVHEI’s items (n = 246).

Scale	Items	Cronbach’s α
Expectancy for Success	EFS1, EFS2, EFS3 and EFS4	.94
Attainment Value	AV1, AV2, AV3 and AV4	.93
Utility Value	UV1, UV2, UV3 and UV4	.86
Intrinsic Value	IV1, IV2, IV3 and IV4	.88
Cost	CST1, CST2, CST3 and CST4	.54
Overall Scale with all the 20 items (Version 1)		.87
Overall Scale with 16 items (Version 2, with CST1 to CST4 excluded)		.91

### Confirmatory factor analysis

Given that the normality assumption was not met in the data distributions, the input matrices for the CFAs were based on Spearman rank correlations, which can accommodate various data distortions, including problems with outliers and non-normality (Coughlan et al., 2007; de Winter et al., 2016). The path diagrams for Version 1 (five factors) and Version 2 (four factors) of the EVHEI and path coefficients from the CFA results are depicted in Figures 1 and 2, respectively. Table 8 provides goodness-of-fit indices used in evaluating the acceptability of the factor solutions for each version.

In general, CFI, NFI and NNFI values between .90 and .95 were deemed to indicate acceptable fit by Brown (2006), while Ab Hamid (2013) deemed value of .90 and above to indicate good fit for these measures. In the present study, comparative fit indices of CFI, NFI and NNFI ranging from .87 to .91 were obtained for the five-factor version, and from .91 to .94 for the four-factor version. While both versions approximated acceptable fit based on these indices, it can be noted that the four-factor version fared slightly better than did the five-factor version, suggesting that again, the fit of the model for Version 2 (i.e., without the four CST items) was superior.

As the  $\chi^2$  test can be affected by sample size (Brown, 2006; Mîndrilă, 2010), the ratio of  $\chi^2$  to degrees of freedom is typically recommended instead. Schreiber et al. (2006) and Mîndrilă (2010) recommended  $\chi^2/df < 3$  as the cutoff. For both versions, the values of  $\chi^2/df$  were below 3 (see Table 8), indicating that both attained acceptable fit based on this criterion. The RMSEA for both versions was also .08, which again suggests an acceptable level of fit based on the recommendations of Browne and Cudeck (1992), Schreiber et al. (2006) and Mîndrilă (2010).

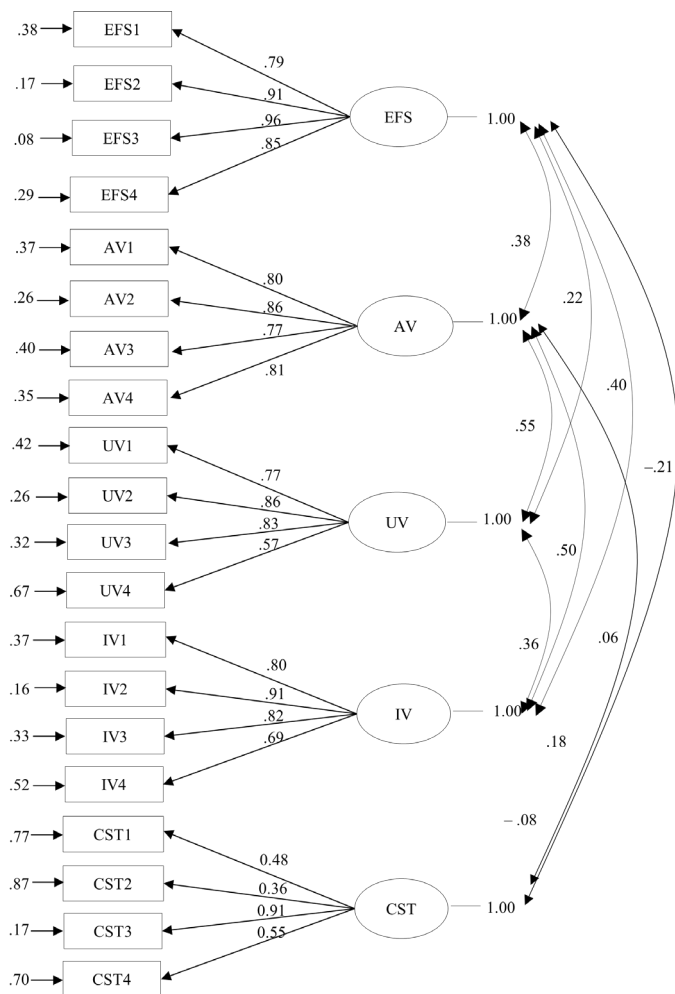


Figure 1. CFA results: Path diagram for Version 1 (Five-factor version).

Table 8. Fit indices for the One-factor, Two-factor, Four-factor and Five-Factor Models.

Fit Index	Based on 20 items			Based on 16 items (CST items excluded)		
	One-factor Model	Two-factor Model	Model 1 (Five factors)	One-factor Model	Two-factor Model	Model 2 (Four factors)
$\chi^2$	$\chi^2(170) = 2,046.38$ ( $p = .00$ )	$\chi^2(169) = 1,303.45$ ( $p = .00$ )	$\chi^2(160) = 442.08$ ( $p = .00$ )	$\chi^2(104) = 1,691.89$ ( $p = .00$ )	$\chi^2(103) = 947.17$ ( $p = .00$ )	$\chi^2(98) = 272.16$ ( $p = .00$ )
$\chi^2/df$	12.04	7.71	2.76	16.27	9.20	2.78
Standardised Root Mean Residual (SRMR)	.15	.13	.07	.16	.12	.06
Root Mean Square Error of Approximation (RMSEA)	.20	.16	.08	.24	.17	.08
Comparative Fit Index (CFI)	.42	.65	.91	.46	.72	.94
Normed Fit Index (NFI)	.41	.62	.87	.45	.69	.91
Non-Normed Fit Index (NNFI)	.36	.61	.90	.38	.67	.93
Goodness-of-Fit Index (GFI)	.51	.64	.86	.51	.66	.89
Adjusted GFI (AGFI)	.40	.55	.82	.36	.55	.85
$\Delta\chi^2$ as compared to one-factor model	-	-	$\Delta\chi^2(10) = 1,604.30$ ( $p < .01$ )	-	-	$\Delta\chi^2(6) = 1,419.73$ ( $p < .01$ )
$\Delta\chi^2$ as compared to two-factor model	-	-	$\Delta\chi^2(9) = 861.37$ ( $p < .01$ )	-	-	$\Delta\chi^2(5) = 675.01$ ( $p < .01$ )

To further assess model adequacy, within both the 20-item and the 16-item versions of the instrument, one-factor and two-factor nested alternative models were also tested. In the case of the 20-item version, the five-factor model derived from the EFA was compared to a one-factor model (with all 20 items loaded on a single factor) and a two-factor model (with the EFS items loaded on one factor and the AV, UV, IV and CST items loaded on the other). The same comparison bases were used for the 16-item version, including a one-factor solution (with all 16 items loaded on a single factor)

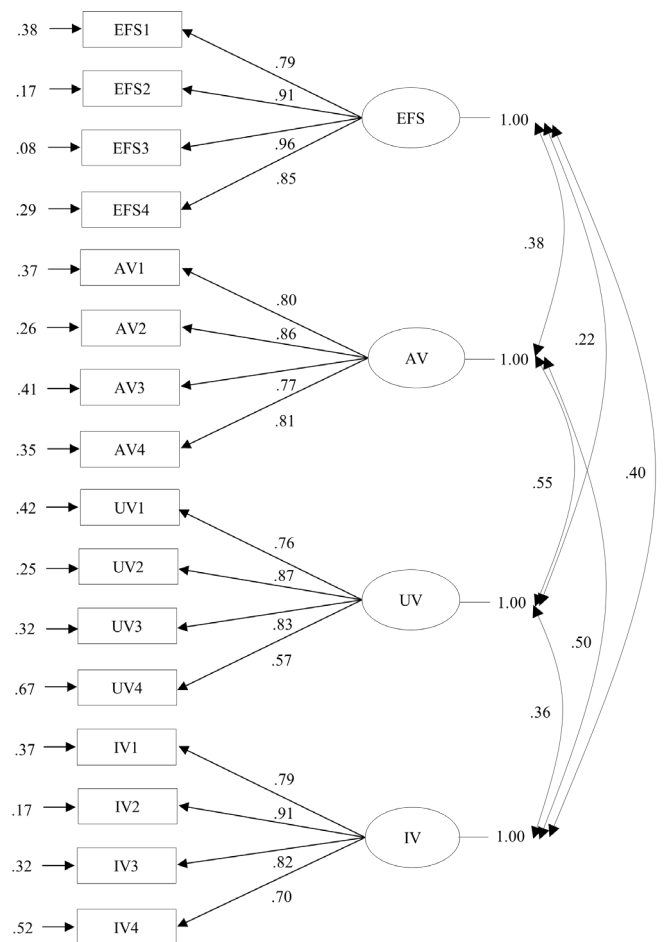


Figure 2. CFA results: Path diagram for Version 2 (Four-factor version).

and a two-factor solution (with the EFS items loaded on one factor and the AV, UV and IV items loaded on another). The change in  $\chi^2$  comparing the EFA models for each version with their respective one- and two-factor models (see last two columns of Table 8) indicated that these models produced significant fit improvement over the one- and the two-factor solutions. Comparisons based on other fit indices (RMSEA, SRMS, CFI, NFI and NNFI) were consistent with this conclusion.

The internal consistency of the items was also evaluated using the CFA subsample using Cronbach's  $\alpha$  coefficients (see Table 9). The Cronbach's  $\alpha$  coefficients for the EFS, AV, UV and IV subscales were all above 0.80. Again, however, the Cronbach's  $\alpha$  coefficient for the CST subscale was noticeably lower than this threshold (.67). This confirmed the relatively weak internal consistency of the CST subscale, which is confirmed further by the fact that the Cronbach's  $\alpha$  coefficient for the overall scale with the four CST items (.85) was lower than for the 16-item version (.90).

Table 9. Cronbach's  $\alpha$  of the EVHEI's items (n = 277).

Scale	Items	Cronbach's $\alpha$
Expectancy for Success	EFS1, EFS2, EFS3 and EFS4	.93
Attainment Value	AV1, AV2, AV3 and AV4	.93
Utility Value	UV1, UV2, UV3 and UV4	.84
Intrinsic Value	IV1, IV2, IV3 and IV4	.88
Cost	CST1, CST2, CST3 and CST4	.67
Overall Scale with all the 20 items (Version 1)		.85
Overall Scale with 16 items (Version 2, with CST1 to CST4 excluded)		.90

## Discussion

The main aim of this study was to develop a stand-alone instrument to measure two key constructs within the EVT framework – expectancies for success and subjective task values constructs – with specific reference to academic success in HE. EFAs and CFAs were conducted to evaluate both a 20-item version (EFS, AV, UI, IV and CST) and a 16-item version (EFS, AV, UI and IV) of the EVHEI, the latter excluding the construct of cost dimension. The ambiguous role of the cost dimension in other literature provided the rationale for the creation and evaluation of these two versions.

With 20 items created to measure the five constructs within the EVT framework (expectancies for success, attainment value, utility value, intrinsic value and cost), the EFA extracted five factors. Using a different dataset, the CFA subsequently validated the suitability of this five-factor structure of the 20-item version of the instrument. In the case of the 16-item version, the EFA extracted four factors, and the subsequent CFA validated this structure.

While acceptable fits were obtained for both versions, the 16-item version generally fared better in this respect, in both the EFAs and CFAs. For example, the four-factor version obtained higher NFI and NNFI indices than did the 20-item version. This was due in large part to the relatively low loadings for specific items within the cost (CST) factor, with corresponding reduced internal consistencies. This suggests some misalignment within the cost factor variable. This aligns with evidence from the literature, which generally underscores the ambiguity of cost as a component of the overall STV construct. Despite this, the fit indices obtained for the 20-item version all fell within or marginally below the acceptable range. Thus, the EVHEI can be used either as a 20- or as a 16-item version. Given the ambiguous role of cost in other papers, this would provide researchers with some flexibility in how they choose to measure the construct of STV in their own studies.

The authors pose that the EVHEI represents a vital development in furthering the potential use of the EVT model in HE. As noted by several authors in the field, the EVT constructs are defined somewhat broadly in theoretical definitions and are inherently linked to a wide array of factors (Wigfield et al., 2009). In light of this, it is unsurprising to find that vastly different measures have been used in different studies to measure the expectancy and task value constructs, as mooted earlier. With appropriate construct measurement using the EVHEI, more meaningful and precise investigations relating motivation and HE success can then be undertaken.

That said, the EVHEI may not be suited for use in all contexts over time. In particular, the HE sector is currently in a state of flux, which also means that students and their motives for engaging in learning will also be so. As such, the kinds of factors which feed the expectancy and task value beliefs formed by students may shift over time. Other factors may also arise as relevant over such time.

With the acceleration of e-learning in HE, which has changed not only the way that students learn and perform academic tasks but also how their performance is assessed, students

are now being given greater autonomy in learning, allowing them to decide not only where learning can take place, but also the pace in which it occurs (Jansen et al., 2017). It is possible that such shifts will also change the most relevant constructs for estimating the level of motivation that students will have in their studies. In such an event, however, the EVHEI could be used as a base framework for developing subsequent instruments. That is, the items can be modified in minor ways to suit different contexts.

There are other aspects that can be taken into consideration for the future development of the instrument. As the participants involved in this study were sampled from a specific institution and had very specific attributes in common (e.g., all were full-time students enrolled in undergraduate programmes offered by a Singapore university), their profiles could also be seen as relatively homogenous. The EVHEI, therefore, would need to be validated further using participants with different profiles to assess the generality of its psychometric properties across populations and contexts.

Although the two versions of the EVHEI instrument provide flexibility in measuring the cost dimension of EVT, this does not imply that the authors themselves consider the cost factor to have an insignificant role in the measurement of motivation. Rather, it is possible that cost should be treated as an independent motivational construct within the EVT model, in line with the Expectancy-Value-Cost model discussed earlier. Such theoretical discussion falls beyond the scope of the current paper. The dual versions of the EVHEI, however, provide researchers with the flexibility to choose whether and how to incorporate cost in the measurement of subjective task values in EVT, depending on the contexts in which they operate.

## References

- Ab Hamid, M. R., Mustafa, Z., Suradi, N. R. M., Idris, F., & Abdullah, M. (2013). Value-based performance excellence measurement for higher education institution: Instrument validation. *Quality & Quantity*, 47(6), 3019–3030. <https://doi.org/10.1007/s11135-012-9699-y>
- Allen, D. (1999). Desire to finish college: An empirical link between motivation and persistence. *Research in Higher Education*, 40(4), 461–485. <https://doi.org/10.1023/A:1018740226006>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman. <https://doi.org/10.1037/10522-094>
- Barron, K. E., & Hulleman, C. S. (2015). Expectancy-value-cost model of motivation. In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (2nd ed., Vol. 8, pp. 503–509). Oxford: Elsevier. <http://dx.doi.org/10.1016/B978-0-08-097086-8.26099-6>
- Battle, A., & Wigfield, A. (2003). College women's value orientations toward family, career, and graduate school. *Journal of Vocational Behavior*, 62(1), 56–75. [https://doi.org/10.1016/S0001-8791\(02\)00037-4](https://doi.org/10.1016/S0001-8791(02)00037-4)

- Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology, 26*(4), 553–570. <https://doi.org/10.1006/ceps.2000.1048>
- Brown, T. (2006). *Confirmatory factor analysis for applied research* (1st Ed.). Guilford Publications. <https://www.guilford.com/books/Confirmatory-Factor-Analysis-for-Applied-Research/Timothy-Brown/9781462515363>
- Browne, M., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230–258. <https://doi.org/10.1177/0049124192021002005>
- Brunhaver, S.R., Carrico, C., Matusovich, H. M., Sama, M., Abhyankar, R., Streveler, R. A., & Sheppard, S. (2017). Measuring students' subjective task values related to the post-undergraduate career search [Paper presentation]. *2017 ASEE (American Society for Engineering Education) Annual Conference & Exposition, Columbus, Ohio*. <https://www.semanticscholar.org/paper/Measuring-students'-subjective-task-values-related-Brunhaver-Carrico/1b5dcc2c42071054726bba54996e0dfac062ad9e>
- Chirinos, D. S. (2017). *Examining dimensions of expectancy-value theory as predictors of U.S. Latino high school students' academic behaviors and mathematics performance* (Order No. 10277657). ProQuest Dissertations & Theses Global. [https://www.researchgate.net/publication/319130449\\_Examining\\_dimensions\\_of\\_expectancy-value\\_theory\\_as\\_predictors\\_of\\_us\\_latino\\_high\\_school\\_students'\\_academic\\_behaviors\\_and\\_mathematics\\_performance](https://www.researchgate.net/publication/319130449_Examining_dimensions_of_expectancy-value_theory_as_predictors_of_us_latino_high_school_students'_academic_behaviors_and_mathematics_performance)
- Coughlan, J., Shale, E., & Dyson, R. (2007). *The effect of different inputs to factor analysis: An example using service quality in UK branch banking*. <https://arrow.tudublin.ie/cgi/viewcontent.cgi?article=1001&context=buschmarcon>
- Credé, M., & Kuncel, N. R. (2008). Study habits, skills and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425–453. <https://doi.org/10.1111/j.1745-6924.2008.00089.x>
- de Winter, J. C. F., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological Methods, 21*(3), 273–290. <https://doi.org/10.1037/met0000079>
- Demetriou, C., & Schmitz-Schiborski, A. (2011). Integration, motivation, strengths and optimism: Retention theories, past, present and future. In R. Hayes (Ed.), *Proceedings of the 7th national symposium on student retention, 2011, Charleston* (pp. 300–312). Norman, OK: The University of Oklahoma. [https://www.researchgate.net/publication/364309350\\_Integration\\_Motivation\\_Strengths\\_and\\_Optimism\\_Retention\\_Theories\\_Past\\_Present\\_and\\_Future\\_Integration\\_Motivation\\_Strengths\\_and\\_Optimism\\_Retention\\_Theories\\_Past\\_Present\\_and\\_Future](https://www.researchgate.net/publication/364309350_Integration_Motivation_Strengths_and_Optimism_Retention_Theories_Past_Present_and_Future_Integration_Motivation_Strengths_and_Optimism_Retention_Theories_Past_Present_and_Future)
- Dietrich, J., Viljaranta, J., Moeller, J., & Kracke, B. (2017). Situational expectancies and task values: Associations with students' effort. *Learning and Instruction, 47*, 53–64. <https://doi.org/10.1016/j.learninstruc.2016.10.009>
- Doménech-Betoret, F. (2006). Testing an instructional model in a university educational setting from the student's perspective. *Learning and Instruction, 16*(5), 450–466. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2006.09.005>
- Doménech F. (2012). *Psicología Educativa: Su Aplicación al Contexto de la Clase [Educational Psychology: its Application in the Classroom Context]*, 13. Castellón: Publicaciones de la Universitat Jaume I. Col·lecció Psique.
- Doménech F. (2013). An instructional model for guiding reflection and research in the classroom: The educational situation quality model. *Electronic Journal of Research in Educational Psychology, 11*, 239–260.
- Doménech-Betoret, F., Abellán-Roselló, L., & Gómez-Artiga, A. (2017). Self-efficacy, satisfaction, and academic achievement: The mediator role of students' expectancy-value beliefs. *Frontiers in Psychology, 8*, 1–12. <https://doi.org/10.3389/fpsyg.2017.01193>
- Eccles, J. S. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives*. San Francisco: Freeman.
- Fabrigar, L., & Wegener, D. (2011). *Exploratory factor analysis* [Adobe Digital Editions]. ProQuest Ebook Central. <https://academic.oup.com/book/41736>
- Fabrigar, L., Wegener, D., Maccallum, R., Strahan, E. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Flake, J., Barron, K., Hulleman, C., McCoach, B., & Welsh, M. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary Educational Psychology, 41*, 232–244. <https://doi.org/10.1016/j.cedpsych.2015.03.002>
- Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior, 26*(4), 331–362. <https://doi.org/10.1002/job.322>
- Gaspard, H., Dicke, A. -L., Flunger, B., Schreier, B., Häfner, I., Trautwein, U., & Nagengast, B. (2015). More value through greater differentiation: Gender differences in value beliefs about math. *Journal of Educational Psychology, 107*(3), 663–677. <https://doi.org/10.1037/edu0000003>
- Gorges, J. (2015). Out of school, out of mind? An expectancy-value analysis of adult learners' motivation. *Journal of Cognitive Education and Psychology, 14*(2), 263–264. <https://www.proquest.com/docview/1685862469?sourcetype=Scholarly%20Journals>
- Gorges, J., & Göke, T. (2015). How do I know what I can do? Anticipating expectancy of success regarding novel academic tasks. *British Journal of Educational Psychology, 85*, 75–90. <https://doi.org/10.1111/bjep.12064>

- Gorges, J., & Kandler, C. (2012). Adults' learning motivation: Expectancy of success, value, and the role of affective memories. *Learning and Individual Differences, 22*(5), 610–617. <https://doi.org/10.1016/j.lindif.2011.09.016>
- Griffin, R., MacKewn, A., Moser, E., & VanVuren, K. W. (2012). Do learning and study skills affect academic performance? An empirical investigation. *Contemporary Issues in Education Research, 5*(2), 109–116. <https://doi.org/10.19030/cier.v5i2.6928>
- Heckman, J. (2007). *Importance of early childhood development: Invest in the very young. Encyclopedia on early childhood development* (2nd ed.). <http://www.child-encyclopedia.com/importance-early-childhood-development/according-experts/invest-very-young>
- Ho, R. (2014). *Handbook of univariate and multivariate data analysis with IBM SPSS* (2nd ed.). Taylor and Francis. <https://doi.org/10.1201/b15605>
- Jansen, R., van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education, 29*(1), 6–27. <https://doi.org/10.1007/s12528-016-9125-x>
- Jones, B. D., Paretti, M. C., Hein, S. F., & Knott, T. W. (2010). An analysis of motivation constructs with first-year engineering students: Relationships among expectancies, values, achievement, and career plans. *Journal of Engineering Education, 99*(4), 319–336. <https://doi.org/10.1002/j.2168-9830.2010.tb01066.x>
- Kusurkar, R. A., Ten Cate, T. J., Vos, C. M. P., Westers, P., & Croiset, G. (2013). How motivation affects academic performance: A structural equation modelling analysis. *Advances in Health Sciences Education, 18*(1), 57–69. <http://dx.doi.org/10.1007/s10459-012-9354-3>
- Lai, E. R. (2011). *Motivation: A literature review research report*. <https://viciongroup.info/assets/files/25.-Motivacin.unarevisindeliteraturalnglsautorEmilyR.Lai.pdf>
- Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don't's, and how-to's. *International Journal of Psychological Research, 3*(1), 97–110. <https://www.semanticscholar.org/paper/How-to-factor-analyze-your-data-right%3A-do%E2%80%99s-and-Matsunaga/1535fea223fabfe05a226b2d76a4b40524710f97>
- Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society (IJDS), 1*, 60–66. <https://doi.org/10.20533/ijds.2040.2570.2010.0010>
- Morrow, J. A., & Ackermann, M. E. (2012). Intention to persist and retention of first-year students: The importance of motivation and sense of belonging. *College Student Journal, 46*(3), 483–491. <https://psycnet.apa.org/record/2012-24487-003>
- Osborne, J. W. (2014). *Best practices in exploratory factor analysis*. CreateSpace Independent Publishing. [https://www.researchgate.net/publication/265248967\\_Best\\_Practices\\_in\\_Exploratory\\_Factor\\_Analysis](https://www.researchgate.net/publication/265248967_Best_Practices_in_Exploratory_Factor_Analysis)
- Post, T., & Walma van Der Molen, J. H. (2019). Development and validation of a questionnaire to measure primary school children's images of and attitudes towards curiosity (The CIAC questionnaire). *Motivation and Emotion, 43*(1), 159–178. <https://doi.org/10.1007/s11031-018-9728-9>
- Robbins, S. B., Lauver, K., Le, H., Davis, D., & Langley, R. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261–288. <https://doi.org/10.1037/0033-2909.130.2.261>
- Roeser, R. W., Midgley, C., & Urdan, T. C. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology, 88*(3), 408–422. <https://doi.org/10.1037/0022-0663.88.3.408>
- Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Broda, M., Spicer, J., Bruner, J., Moeller, J., Linnansaari, J., Juuti, K., & Viljaranta, J. (2016). Investigating optimal learning moments in U.S. and Finnish science classes. *Journal of Research in Science Teaching, 53*, 400–421. <https://doi.org/10.1002/tea.21306>
- Schreiber, J., Nora, A., Stage, F., Barlow, E., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323–338. <https://doi.org/10.3200/JOER.99.6.323-338>
- Sogunro, O. A. (2015). Motivating factors for adult learners in higher education. *International Journal of Higher Education, 4*(1), 22–37. <https://doi.org/10.5430/ijhe.v4n1p22>
- Steinmayr, R., & Spinath, B. (2009). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences, 19*, 80–90. <https://doi.org/10.1016/j.lindif.2008.05.004>
- Tabachnick, B., & Fidell, L. (2013). *Using multivariate statistics* (6th ed.). Pearson Education.
- Terenzini, P. T., & Reason, R. D. (2005, November). Parsing the first year of college: A conceptual framework for studying college impact [Paper presentation]. *The annual meeting of the association for the study of higher education*, Philadelphia, PA. <https://docplayer.net/15292255-Parsing-the-first-year-of-college-a-conceptual-framework-for-studying-college-impacts.html>
- Tinto, V., & Pusser, B. (2006, November). Moving from theory to action: Building a model of institutional action for student success [Paper presentation]. *The national symposium on postsecondary student success*, Washington, DC. [https://nces.ed.gov/npec/pdf/Tinto\\_Pusser\\_Report.pdf](https://nces.ed.gov/npec/pdf/Tinto_Pusser_Report.pdf)
- van der Putten, S. A. (2017). Are motivational theories too

general to be applied in education? *SFU Ed Review Special Issue 2017*, 1–12. <https://journals.lib.sfu.ca/index.php/sfuer/article/download/311/174/>

VanZile-Tamsen, C. (2001). The predictive power of expectancy of success and task value for college students' self-regulated strategy use. *Journal of College Student Development*, 42(3), 233–241. [https://www.researchgate.net/publication/234573795\\_The\\_Predictive\\_Power\\_of\\_Expectancy\\_of\\_Success\\_and\\_Task\\_Value\\_for\\_College\\_Students'\\_Self-Regulated\\_Strategy\\_Use](https://www.researchgate.net/publication/234573795_The_Predictive_Power_of_Expectancy_of_Success_and_Task_Value_for_College_Students'_Self-Regulated_Strategy_Use)

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81. <https://doi.org/10.1006/ceps.1999.1015>

Wigfield, A., Tonks, S., & Klauda, S. L. (2009). Expectancy-value theory. In K. R. Wentzel, & D. B. Miele (Eds.), *Handbook of motivation in school* (pp. 55–75). <https://psycnet.apa.org/record/2009-24219-004>

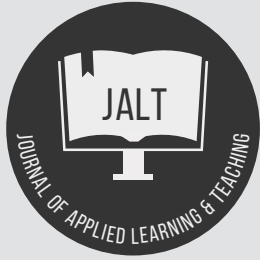
Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79–94. <https://doaj.org/article/2687c98d82474a6d8ccac214e7f31288>

Yoo, S., & Huang, W. (2013). Engaging online adult learners in higher education: Motivational factors impacted by gender, age, and prior experiences. *Journal of Continuing Higher Education*, 61(3), 151–164. <https://doi.org/10.1080/07377363.2013.836823>

Zhou, J., & Urhahne, D. (2013). Teacher judgment, student motivation, and the mediating effect of attributions. *European Journal of Psychology of Education*, 28(2), 275–295. <https://doi.org/10.1007/s10212-012-0114-9>

Zimmermann, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, 29(3). <https://doi.org/10.3102/00028312029003663>

Copyright: © 2024. Wan Hoong Wong and Elaine Chapman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Advice from retired secondary school principals in Ireland on how to lead as a principal

---

Finn Ó Murchú<sup>A</sup>

A

*Associate Professor and Head of School at MIC School of Education, Thurles, Ireland*

Tom O'Donoghue<sup>B</sup>

B

*Emeritus Professor, Graduate School University of Western Australia, Perth*

---

### Keywords

Ireland;  
leaders;  
principal;  
secondary schools.

---

### Abstract

At the core of this paper is a report on a study aimed at generating theory about the perspectives of recently retired secondary school principals in Ireland (n=15) on how beginning secondary principals in the nation should lead. The hope from the outset was that the result could be drawn upon to offer supporting insights to providers, including university-led programmes, so as to inform the pre-service and ongoing preparation of aspiring and appointed principals. What follows is presented in four parts. First, an exposition is provided on the rationale that informed that aim. Secondly, recent developments in relation to school leadership in Ireland are considered. Thirdly, details on the conduct of the study are presented. Fourthly, the study results are outlined.

---

### Correspondence

Finn.omurchu@mic.ul.ie <sup>A</sup>

---

### Article Info

Received 11 September 2023

Received in revised form 9 December 2023

Accepted 4 January 2024

Available online 15 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.6>



## Rationale for the study

We were explicit from the outset that we subscribe to a view that while the most influential model of school leadership for many years has been focused on authority, power, structures, job descriptions, targets and performance management, it should no longer be the sole one advocated by policymakers. That, we hold, is partly because in a complex and multi-layered world, one should no longer accept the conventional idea of 'good' leadership – leadership that promises more success and less failure – as being the result of the effort of a single individual.

At the same time, we hold to the view that beginning school principals need to have a good grasp of the fundamental tenets of the range of leadership types, including managerial leadership (Henkel et al., 2019), instructional leadership (Hallinger et al., 2020), transactional leadership/transformational leadership (Purwanto, 2020; Da'as, 2023), moral leadership (Wise & Slater, 2020), invitational leadership (Egley, 2003), interpersonal leadership (Lamm et al., 2016) and distributed leadership (Spillane, 2012), amongst others. However, we are also clear on the necessity of understanding people's contextual realities before introducing changes aimed at improving the quality of education in any context. On that, we have long been influenced by the argument of Fullan (1982, p. 149) that "in order to effect improvement, that is, to effect an introduced change which has the promise of increasing success and decreasing failure, the world of the people most closely involved must be understood". He went on to say that those who are trying to promote change – amongst whom we include beginning school principals – there is great benefit to be gained from reflecting on interpretivist studies of the leadership views of leaders. This, he argued, is because to have 'good' leadership, the world of the people most closely involved in the education enterprise must be understood.

Around the same time, Hargreaves also held that engaging in interpretivist studies in order to construct theory is important. To contend thus, he argued, is to hold that generating concepts, categories, propositions, models, typologies, and the relationships between all five of these, can result in us having "a language for speaking about that which is not normally spoken about: the ineffable is rendered articulate" (Hargreaves, 1993). Professionals can then be introduced to that 'language', he went on, such that it can act as a mirror that "reflects man [sic] back to himself", thus providing one with "an opportunity to judge and appraise the reflection" one sees and, perhaps as a result, prompting one to seek to change oneself and one's professional world (p. 149).

We are clear too regarding the folly of those who expect that clear directives for practice can be deduced from theory generated. Over fifty years ago Entwistle argued that there never "can be a one-to-one relationship between theory and practice if by this we mean theory which predicts every contingency in a practical situation. A theory gains its relevance to every conceivable situation by being an exact account of none of them.... The fault for the theory-practice gap may lie not in the theory but in the unrealistic expectations of practitioners" (Entwistle,

1971, p. 98). Over ten years later, Eisner (1984) argued that due to the changing uniqueness of the practical situations that make up the domain of education, only a portion of professional practice can be approached as a prescriptive science. The gap between general prescriptive frameworks and successful practice is, he held, dependent more on the reflective intuition, the craft, and the art of the professional practitioner than on a prescriptive theory, method, or model.

Informed by the positions outlined above, then, we take the view that the results of studies that lead to the generation of theory generated within a particular situation, or situations, or with a particular group of individuals, can be informative for the creation of professional development programmes. This is because they can be a source to assist in breaking away from a notion of continuing professional development as being concerned only with instrumental ends achievable through, "the recipes of tried and true practices legitimated by unexamined experiences or uncritically accepted research findings" and towards one of "developing reflective practitioners who are able to understand, challenge and transform" (Sachs & Logan, 1990, p. 479). In the language of Stenhouse (1975), this is a view that such studies can aid in the development of the capacity of educationists to understand relationships and make judgements by constituting frameworks for others within which they can think. Relatedly, it is a view that no claim can be made for the 'generalisability' of interpretivist theories generated in one situation to another. Instead, as Stake (1978) put it, theory generated through interpretive studies undertaken with small populations may be in harmony with the reader's experience and thus a natural basis for generalisation.

The view outlined above is that readers may be able to relate to a study and perhaps gain an understanding from it of their own and others' situations. As Kennedy (1979), in the same vein, put it, generalisability is ultimately related to what the reader is trying to learn from such studies. Strauss and Corbin (1994, p. 279) put it in similar terms when they stated that "those examining an interpretivist theory need to consider the extent to which their own situation fits the theory", along with how it might fit and how it might not. The demand, thus, of those engaged in professional development programmes in which theory is presented and where there is an acceptance that it consists of systematic statements of plausible relationships, is to have an openness based on the 'forever' provisional character of every educationist.

In relation to the latter matter, Lincoln and Guba's (1985) notion of 'transferability' is instructive. They contended that those who generate interpretivist theory cannot specify the external validity of an inquiry. Rather, alongside the theoretical constructs generated, he or she can provide "the thick description necessary to enable someone interested in making a transfer to reach a conclusion about whether the transfer can be contemplated as a possibility" (Lincoln & Guba, 1985, p. 316). Uhrmacher (1993) argued along similar lines that one can provide the reader with an understanding of the major themes that run through the cases under study. In turn, these themes can provide one with theories or guides for anticipating what may be found in other situations. To that, we add that when used in professional development programmes they can also suggest practices

to the participants that are worth putting to the test of practice to see if they lead to 'good' leadership on their part within their particular circumstances. We also hold that such suggestions from beginning school leaders are likely to be forthcoming when it is clear to them that what is being presented for cogitation is embedded in the realities of workplaces found in schools and in the environments in which they are located.

### **Recent developments in relation to school leadership preparation in Ireland**

The study was conducted with recently retired school principals in the Republic of Ireland and is likely to have its greatest utility when used in such a setting for professional development programmes with aspiring and newly appointed principals. That said, there is no reason why it would not also be of some value to their peers in other countries. After all, it has been known for a long time that making comparisons and contrasts with one's own situation can be of great benefit in assisting one in reflecting on one's own situation. Both points, in turn, mean that there is a need to provide an exposition like that which follows on from the context within which studies of the type being promoted were conducted.

Across the 15 participants' timeframe of principalship, 1982-2018, policy expectations associated with the roles and responsibilities of school principals have evolved significantly, gaining considerable pace in more recent times. Of note is a movement that commenced by determining an understanding of what principals need to do, to an understanding of what supports are needed to enhance principals in their work. The origins of that movement were inspired by the authors of a 1991 OECD report on Ireland, a number of whose recommendations were captured in the Education Act of 1998. Relatedly, the following list of official responsibilities were required of principals; encourage and foster learning in students; regularly evaluate students and periodically report to the students and their parents, promote co-operation between the school and the community; be responsible for the day-to-day management of the school, including guidance and direction of the teachers and other staff of the school, and be accountable to the board for that management; provide leadership to the teachers and other staff and the students of the school; be responsible for the creation, together with the board, parents of students and the teachers, of a school environment which is supportive of learning among the students and which promotes the professional development of the teachers; set objectives for the school (under the direction of the board and, in consultation with the teachers, the parents and, to the extent appropriate to their age and experience, the students) and monitor their achievement; and encourage the involvement of parents of students in the school in the education of those students and in the achievement of the objectives of the school.

Centrally organised supports for school leadership soon followed and were made widely available in the first decade of the century with particular emphasis placed on distributed and instructional styles of leadership, albeit without any clear

and detailed explanation of what these concepts might mean at the national or local level. School Development Planning Initiative (SDPI) initiatives were aimed at improving school practice in relation to a series of indicators which informed the publication of *Looking at our schools* (Department of Education and Science, Ireland, 2022). The later document focuses in particular on 'Teaching and Learning' and 'Leadership and Management' and informs not only the work of schools but also school inspections and the criteria used for appointments to positions of middle and senior leadership. Of note, the most recent edition of *Looking at our schools* (Department of Education and Science, Ireland, 2022) more explicitly details the role of school leadership and its connection to both school self-evaluation (SSE) and the criteria for leadership appointments.

Concurrently, work on engaging in internal school reviews and external school evaluations in search of school improvement also came to be emphasised and was supported at post-primary school level by a growing number of Department of Education (DE) schools' inspectors. In turn, those at the centre in the national Department of Education (DE) also lent their support to the establishment of the National Association of Principals and Deputy Principals (NAPD) in 1998 and the provision of support in the form of a suite of leadership development programmes for newly appointed principals. In more recent times, the latter has been supplemented by the work of a new Centre for School Leadership (CSL), established in 2015. Professionals at this centre offer guidance and support to those occupying different positions across the leadership strata in schools. Furthermore, largely through retired school principals they specifically offer coaching and mentoring support to newly appointed principals and to principals who feel that such may be of assistance to them at any particular time. In addition, they offer endorsement of a myriad of leadership programmes, provided by higher education institutions in order to encourage aspirant and appointed leaders to obtain accredited qualifications from a wide range of providers. That said, as highlighted by Murphy (2023):

Notably, formal leadership preparation for senior (the principalship or deputy principalship) or middle leadership (assistant principals, of which there are two categories, Assistant Principal 1 and Assistant Principal 2) is not, as yet, officially required (p. 30).

While not, as of yet, formally required to be appointed to leadership roles in Irish secondary schools, it is clear from Murphy's insightful research (2023) that it is now accepted by aspiring leaders that it is 'impossible to secure a senior leadership role without engaging in formal leadership preparation' (p. 30). However, the appetite to become or remain a school leader would appear to have waned in recent times (Clemens et al., 2016; National Association of Principals and Deputy Principals, 2020). These two context-sensitive scenarios are interconnected and in such a context, the advice, in our small study, from recently retired principals may prove of value to those seeking to support and to design professional learning opportunities for aspiring and appointed school leaders.

Overall, a shift has taken place towards greater coherence in relation to school leadership and a broader interpretation of school leadership than previously held in Ireland. Middle leaders and those without formal roles (described as teacher leaders) have been part of that shift, especially since 2018. Such is captured with a renewed policy emphasis and entreaties associated with promoting both distributed and instructional leadership (Department of Education, 2018) among school personnel. The promotion of collaborative leadership practices within and across schools continues to be a policy imperative heightened more recently by responses to the Covid pandemic. Nevertheless, recruitment and retention of school principals continues to be a challenge. Thus, two days before the Covid-19 pandemic first closed schools in Ireland, the NAPD stated (Irish Times, March 10th 2020) that it was critical to address potential solutions and support, so as to ensure the position of school principals continued to attract the best talent and expertise into the future. Adding, that research from their own survey found that less than one in three school leaders predicted they would still be in a leadership role in five years (NAPD, 2020).

## The conduct of the study

The study, it will be recalled, aimed at seeking insights to inform the pre-service and on-going preparation of neophytes for the position of second-level school principal in Ireland, was undertaken by the present writer and reported later in this thesis. To that end, the authors set out to generate a theory on the perspectives of recently retired second-level school principals in Ireland on how beginning second-level principals in the nation should lead. The associated central guiding research question was as follows: What is the most comprehensive theory that can be generated on the perspectives of recently retired second-level school principals in Ireland on how beginning second-level principals in the nation should lead. That in turn led to the generation of the following sub-research guiding questions based on the four component parts of 'perspectives' as explicated by Blackledge and Hunt (1985):

1. What intentions do recently retired second-level school principals in Ireland say that beginning second-level principals in the nation should have on how to lead in schools and what reasons do they give for having them?
2. What strategies do retired second-level school principals in Ireland say that beginning second-level principals in the nation should use when leading in schools and what reasons do they give for having them?
3. What significance do recently retired second-level school principals in Ireland say that beginning second-level principals in the nation should attach to their job of leading in schools and what reasons do they give for having them?
4. What outcomes do recently retired second-level school principals in Ireland say that beginning

second-level principals in the nation should expect as a result of how they lead in schools and what reasons do they give for having them?

Grounded theory was the associated research methodology adopted. That was because it makes its greatest contribution in areas like that studied in which little research has been undertaken (Bryant, 2013). Fernandez (in Walsh et al., 2015) put the associated argument well in explaining that grounded theory is an approach to research that "privileges context" over academic theory, and thus is particularly applicable when the impacts of specific cultural contexts are at the heart of a study. Levina (in Walsh et al., 2015) developed it further, stating that grounded theory "embraces the richness and uniqueness of the context without necessarily ignoring the development of theory applicable to other phenomena and contexts" (p. 592).

Consistent with the grounded theory research approach, the researcher selected the research methods of semi-structured interviews with individuals and the analytical approach of open coding. Purposeful participant selection was used to ensure that information-rich participants were interviewed (Patton, 1990). On that, as the study aimed to generate theory not yet developed it was not necessary to work with large numbers of participants (Author 2). It was realised too that collecting too much data could impede analysis and lead to 'conceptual blindness' (Morse, 2010). Therefore, it was held that it was important for participants to be chosen whose responses would be likely to maximise the quality, rather than the quantity of data.

In total, 15 principals were involved in the study. Of those researched, 5 had been appointed to their first position as principal in the 1980s. The other 10 took up their first position as principal just before the turn of the century. All had been retired for less than 3 years. Following being contacted by email and invited to participate in the study, each participant who volunteered received a letter outlining the research.

An interview schedule was developed based on the four guiding questions outlined above. The principals participated in individual interviews. The actual order of the questions was partly determined by the interviewee, with the interviewer allowing him or her to take the lead. The interviews were audio-recorded with the permission of the participant and then transcribed to allow ease of analysis according to procedures outlined by Jamshed (2014).

In accordance with grounded theory approaches to analysis (Punch & Oancea, 2014), the authors interrogated the data to generate conceptual categories, found relationships between them, and conceptualised these relationships. Coding was undertaken at three levels: open coding, axial coding, and selective coding (Corbin & Strauss, 2008). The authors engaged in these processes until they considered they had reached 'saturation'.

## Overview of results

The following 'story-line' (ref.) was generated in relation to the results:

*All of the practices advocated by retired teachers for beginning principals have already been highlighted in the academic literature as being necessary for engaging in successful leadership. Moreover, they do not neglect to highlight any of the practices that have prominence in that body of work. While not offering any new insights in relation to each of them, they express, both explicitly and implicitly, a view that there is a hierarchy in terms of practices to which leaders should attend. On that, there is an emphasis much greater than what one gets in the current literature on the importance of maintaining a physically healthy and psychologically balanced life if one is to be a successful leader, and one that is seen to override all other considerations that one should address when approaching the job.*

*Next in line in terms of a hierarchy of considerations, and again very prominent, is the importance attached to taking context into consideration and showing appreciation before one takes up one's position of what one has inherited. Unlike in relation to other aspects of education, including curriculum and pedagogy, the 'understanding of context' element of this is something that has come to be advocated with great strength in the field of leadership only in recent years. Even then, it is highlighted in their writing by only a handful of scholars.*

*Third in line in terms of a hierarchy of considerations is a clear notion of how having taken steps to ensure one can maintain a physically healthy and psychologically balanced life and to understand the context in which one is operating is that in subsequently adopting leadership practices, one should engage in associated considerations in a particular logically structured manner. Such an approach, it is indicated is necessary if one is to provide leadership that is systematic and effective and is likely to maximise the achievement of positive results. It involves, it is held, first establishing clearly the overall aims one is trying to achieve, then engaging in comprehensive preparation, then putting in place what are termed 'necessary safeguards', then engaging in 'positive action', and then setting out to harness as much positive action as one can to realise one's aims.*

**The three main aspects of this storyline are now elaborated upon.**

### ***Maintaining a balanced life***

Retired principals in the advice they give to neophytes highlight above all else the importance of maintaining a balanced lifestyle at all times. On that, they draw attention to three practices they consider new principals should utilise. They are maintaining a balanced life, maintaining a sense of one's own mortality, and taking steps to not only take pleasure from the achievements of the schools' students, but also from communicating them widely. These matters have been receiving attention in the academic literature, albeit only in recent years and also by only a handful of scholars. Each is now outlined in turn.

### ***Staying physically and mentally fit***

One can only maintain a balanced life, retired principals say, if one is physically and mentally fit. Thus, they advise neophytes to plan and maintain a regular and sensible physical exercise regime. Equally, they stress the need for one to arrange to make an appointment with one's general practitioner twice a year for appropriate blood tests. That, they hold, is important not just in case there are any threats that one may likely to contract a major illness, but also to discuss and get advice on anything that may be resulting in stress and sleep deprivation, sleep issues and stress.

Retired principals equally stress the need for one to monitor what they term one's 'mental fitness'. This is put forward as having both a positive and negative dimension. The former relates to the need, as it is put, for one "to be true to oneself" and to be able to benefit from the satisfaction one can get from positive comments made by members of the school community regarding one's role as a leader when one presents oneself as someone they "see as human". Equally, it is argued, in order to bring enthusiasm and energy to one's work, one needs to maintain regular friendships with non-work-related peers, regularly have holidays, and maintain at least one major non-work-related interest or hobby.

A negative matter raised by retired school principals regarding the disposition that neophytes should cultivate is that one should not take home one's work. By this is meant that one should not take administrative work home to address in after-work hours. That is proposed so that one has an opportunity to refresh oneself in mind and body for the next day's work. It is, however, also proposed so that, as it is put, one does not in the case of those with partners and possibly other family members "burden those at home". For one to do so is seen not only to be unfair to those in one's care on the home front, but also because one would be turning one's back on the possibility of being refreshed through the dynamics of post-interpersonal relationships and thus return to one's workplace emotionally refreshed and in a good state of mind to have a positive influence on others.

### ***Maintaining a sense of one's own mortality***

A second way of maintaining a balanced lifestyle so that one continues to be a positive-minded and effective school leaders, retired principals say, is to ensure one maintains a sense of one's own mortality. A valuable attitude to adopt in that regard, one argued, is to "remind oneself that one is a cog on a wheel and that cogs can wear out". Another, in like manner, stated that it is helpful to adopt a view that "principals only borrow their schools for a limited period of time".

Further, regarding this matter of maintaining a sense of one's own mortality, retired principals state also that one needs to regularly work hard with oneself on staying calm, and especially when under pressure. Regarding the latter, they add, one needs to recognise what many term the "stress indicators and triggers" that could make one deviate from such a demeanour. That, in turn, they contend, means

putting oneself in a situation where one can listen with an open mind to what those under one's authority are saying. As one put it, "once you understand where people are coming from, you may feel less angry towards them and can create a gap within yourself between stimulus and reaction to produce a better response – or to do nothing." "So", she concluded, "mind the gap".

### ***Taking pleasure from students' achievements and broadcasting them***

A third way of maintaining a balanced lifestyle so that one continues to be a positive-minded and effective school leaders, retired principals say, is to take pleasure from students' achievements and broadcasting them. That, it is asserted, should involve attending and enjoying their participation in sporting events and concerts. Moreover, it is added, one should generate means for celebrating their achievements and those of staff. Amongst many approaches advocated in this regard is the creation of a gallery of photos for one's end-of year report in which one should detail everyone and everything that appears within it. All of this, it is implied, will help in the generation and re-generation of one's enthusiasm for one's work as a school leader as a result of both the intrinsic joy, one can experience by acting along such lines and also from the positive feedback one is likely to get from students and staff members.

### ***Paying attention to context***

Next in prominence in the hierarchy of considerations promoted by recently retired school principals for adoption by neophytes is the importance attached to taking context into consideration and making public one's appreciation of it before one takes up one's position. This matter has also been given increased emphasis in the academic literature in recent years, again albeit by only a handful of scholars. Moreover, the notion of recently retired principals that it should take precedence over addressing such other aspects of education as curriculum and pedagogy, has not been a feature of the corpus of academic works in the field. Overall, they also articulate the importance they attach to the position, identify issues that might act as obstacles to putting it into practice, and identify what they see as a set of practices on how to proceed.

### ***The importance they attach to this position***

Three matters are highlighted by retired principals on the importance of paying attention to the context of the school in which one is employed and why one needs to make this known to those who make up the school community. First, they point out that while a newly appointed school leader is likely to have a range of competencies in relation to such matters as teaching and learning, and administration, such skills, as one stated, "must be allied to the wise review and understanding of the socio-emotional context of the organisation". That, in turn, it is asserted, requires that one knows the strengths and weaknesses of the school. Moreover, it is maintained that this requires that one not

only comes to such an understanding with the perspective of an outsider and makes judgments according to one's own criteria but just, if not even more important, in terms of how participants see themselves, the school culture, and why, within it, they say that things are justifiably done as they are.

### ***Perceived potential obstacles***

The second matter highlighted by retired principals on the importance of paying attention to the context of the school in which one is employed is that one should not assume that all schools are the same. On that, they stress how vital it is that one works hard on trying to avoid making comparisons with previous school(s) both in doing one's own assessment and in sharing one's perspectives with staff and community members. A school's culture, as a participant declared, can be like an iceberg in the sense that only a small amount of it is readily visible. To view the rest accurately, she went on, requires that one engages in a lot of difficult and open-minded work. And in doing that, she concluded, one should not assume that one's values are the same as the school's deputy principal and other members of staff, and also that one is always correct if there are differences of view.

### ***Suggested practices on how to proceed***

Thirdly, retired principals suggest useful practices in which one can engage in order to arrive at both outsider and insider perspectives on the context of a school. At a most basic level, it is important, as alleged, that one meets staff members individually and regularly, albeit not in regular situations socially. That could commence by adopting a practice of greeting each every day. Relatedly, it is recommended that from the outset, one needs to show appreciation for what one has inherited. On that, for example, it is likely that the timetable for teaching is in place at the time of one's arrival. A qualifier is added though. One person put that well in proposing that "one should try to avoid remodelling all institutional practices. Brush the cobwebs off the furniture rather than throw it all out".

It is argued too that it is important to develop a habit of showing interest regularly in what staff members do and what competencies they possess, and that that interest in the staff not only be genuine interest but also be seen as such. A start can be made in that regard by having a major involvement in drawing up the school timetable. To that is added that all staff members, including ancillary staff, as well as students, parents, members of the school's board of management, need to be listened to, to know that they are being listened to, and to know that it is done in a non-judgmental manner. The importance of getting to know caretakers is given special mention on the grounds that they will help one to know much about the school buildings, which can influence school performance in many ways and thus need to be managed in the interest of providing leadership aimed at generating positive learning communities. Getting to know the views of other school support staff is also stressed, especially in view of the wealth of institutional knowledge they often possess that one can harness. To that is added a need to know what approach is taken by those

support staff who are the first anyone meets on entry to the school. Having an open-door policy is also promoted as being conducive to fostering a positive learning community in which there are trust and openness, two properties, it is asserted, of good leadership.

Retired principals also allege that, while it can be a mammoth task, it is well worthwhile trying to remember each student's name and addressing him or her wherever possible by his or her first name. A constant refrain is that one should regularly engage with the students in groups in public places during recess period, listen to their stories, and ask them how they are getting on with their schoolwork and in life in general. The value in that practice, it is added, can be amplified if one can teach a number of classes oneself since, among other advantages, it can help in the process, as more than one stated, "of listening closely to the students".

### **Approaching leadership tasks in a logical manner**

Third in line in terms of a hierarchy of considerations is a clear notion of how having taken steps to ensure one can maintain a physically healthy and psychologically balanced life and understanding the context in which one is operating is that in subsequently adopting leadership practices, one should engage in associated considerations in a particular logically structured manner. Such an approach, it is indicated is necessary if one is to provide leadership that is systematic and effective and is likely to maximise the achievement of positive results. It involves, it is held, first establishing clearly the overall aims one is trying to achieve, then engaging in comprehensive preparation, and then engaging in comprehensive planning. Furthermore, they identify in relation to the latter activities a range of practices concerned with putting in place what they view as necessary safeguards and harnessing as much positive action as one can to realise one's aims.

#### **Step one**

A first step outlined by retired principals for approaching what they deem to be the most important tasks associated with the role of a principal is that one needs to clearly establish one's associated overall aims. Indeed, they go on to state what those aims should be. That was put succinctly by one participant:

It is vital that one seeks to build an institutional identity and culture that has the capacity to build and adapt to challenges and uncertainty. Thus, the principal has to try to provide direction and clarity on the one hand, while at the same time seeking to encourage the school individuals and teams to be creative and developmental. It requires that one works to balance the demands of metrics, examinations, and internal and external evaluations with trying to be inspirational.

It is accepted that the task is not an easy one since school life is always very busy. Accordingly, it is concluded, one needs to encourage habits within oneself to ensure that one

never loses sight of such an aim.

A number of associated practices are also recommended by the retired principals. One of those is that one should set aside some time each day to reflect on how one is progressing in relation to the school mission. Some recommend keeping a diary which can be revisited at different times and support reflection and subsequent actions. One should also use such occasions to set related short-term goals. In a similar vein, one should seek out opportunities to regularly clarify and reinforce the mission ideals amongst all members of the school community. That could involve having outlines of the mission in the staff handbook and drawing attention to it at staff meetings, assemblies, open nights, and award ceremonies.

#### **Step two**

The second step outlined by retired principals for approaching what they deem to be the most important tasks associated with the role of a principal is that one should engage in comprehensive preparation aimed at putting in place what they term "necessary safeguards" in relation to "vision destroyers." To that end, it is held, having a 'significant other', with whom one can regularly confide is valuable. The notion is that such an individual should be a mentor located outside the organisation who can offer support and debriefing sessions. Ideally, it is argued, such individuals should be detached and resilient and be able to provide guidance when required without mandating.

Retired principals also argue that it is helpful to cogitate ways in which staff members can operate to destroy a school's vision. One way detailed is "by maintaining a detached disposition involving denying and disconnecting from feelings" and "allowing your school to become emotionally anorexic". Another is by groups colluding and distributing degrees of responsibility and irresponsibility. That, as a former principal voiced it, can involve "locating incompetence and poor-quality teaching away from where it is visible, but not addressing the core issue". Another way yet again is through what is termed "purposeful obscurity". That is defined as acting to avoid locating where responsibility lies". A final and related way identified by retired school principals that staff members can operate to destroy a school's vision is by reducing the impact of responsibility by delegating upwards, to the principal. Moreover, it is added, one is likely to maximise such threats to one's efforts if one is ever "heard talking either within or outside the school about work being a casual affair" even though one may not mean it.

It is also stressed that in anticipating negative actions one needs to be careful in establishing one's relationship with the deputy principal who is already working in the school on one's arrival. Sometimes one might find out he or she is aggrieved at not having won one's post. The issue as one individual voiced it, "the deputy principal is an essential component of your leadership role. Thus, any concerns should be managed in a professional manner". "This person", it is added, "really has the capacity to make your first year smooth or very bumpy". One needs, it is stated, "to always

indicate positive regard for the individual in question” while not “handing over your power and not adopting an attitude that the individual in question knows because he or she was there before you arrived”. Overall, it is concluded, what is required is that one strikes a very delicate balance in relation to the matters detailed and that this can be made possible through having formal daily meetings with the deputy principal and regular ones with those involved in middle management.

It is crucial too, it is held, that one be seen not to identify with any one group within the school community. Accordingly, one should send all information to all staff at the same time via email. In addition, it is recommended that one will have more time to get established and have the energy to address ‘detractors’ if one does not try to do too much in one’s first year or two. Indeed, it is held that one should not attempt any initiatives until mid-way through one’s first year and that that should be only one initiative and one should lead it oneself. On that, more than one person concluded that any changes one envisages should be pursued gradually and be seen to be beneficial for the whole school community. To that end, they added, the principal needs to “anticipate who is going to be affected by a change and consult with them beforehand”.

### **Step three**

The third step outlined by retired principals for approaching what they deem to be most important tasks associated with the role of principal is that one should harness as much positive action as one can to realise one’s aims. For that, it is contended, one needs to prepare by reading regularly on how a school functions while not trying to know everything there is to know, including about the school. That means that one needs to be committed to engaging in inclusive action and delegating responsibility. One also needs, it is added, to adopt and be seen to maintain a positive work ethic. The latter could include being the first to meet deadlines set for various tasks. It can also include being the first to arrive at the school each morning, being available and approachable with staff and students, and being visible through constantly wandering about the school.

Retired principals also say that positive actions one can take to try to realise one’s aims can involve joining school principals’ associations, attending regional and national management body meetings and continuing professional development days, and sharing management body material with the members of the school’s board of management. Indeed, it is added that such actions can impress on the individuals concerned the importance of the principal’s role and that those who embrace that view can be an invaluable resource available when planning each school year in advance.

Having regular meetings of the school’s teaching staff is also recommended. Those, it is held, should be geared towards promoting a collaborative and self-reflective practice that reduces the strain and onus on the individual leader to be the one with all the answers. That can be part of a strategy, it is argued, aimed at building trust amongst oneself and

one’s colleagues. Mutual trust, it is argued, can also grow when teachers are asked for their cooperation. As one retired principal put it, “that can serve as an antidote to the situation whereby teachers have been affirmed as sole traders all of their lives and where working together was often seen as contrived collegiality”. At the same time, it is recognised that sometimes there is value in engaging ‘outside experts’ to facilitate planning days aimed at reviewing challenges. Those could include, for example, giving staff opportunities to be part of core school teams and to benefit from a change of posts.

Core to the practices proposed, it is suggested, is principals communicating regularly with all, including parents and students. Equally, they should be decisive through engaging in critical dialogue and not be afraid to be creative and innovative. Moreover, they should be forgiving of their own mistakes and should not let the fear of making mistakes hinder the school’s teachers in executing their creativity as long as they do so responsibly. Added to that, they should be prepared to admit when they have made mistakes and be not afraid of making apologies.

### **Conclusion**

At the core of this paper is a report on a study aimed at generating theory about the perspectives of recently retired secondary school principals in Ireland on how beginning principals in the nation should lead. The hope from the outset was that the result could be drawn upon to offer supporting research to providers so as to inform the pre-service and ongoing preparation of aspiring and appointed principals. What followed was presented in four parts. First, an exposition is provided on the rationale that informed that aim. Secondly, recent developments in relation to school leadership in Ireland were considered. Thirdly, details on the conduct of the study were presented. Fourthly, the study results were outlined.

The exposition on the results generated, does not contain copious details. Instead, influenced by Weber’s (1994) notion of ‘ideal types’, we set out to produce an account that would be helpful for comparative study by virtue of being a synthesis of “diffuse, discrete, more or less present and occasionally absent concrete individual phenomena which are arranged . . . into a unified analytical construct” (p. 92) In saying this, we recognise that there are retired principals whose perspectives most likely differ from those we foreground. Accordingly, we invite readers to be open to the possibility that the efficacy of our interpretations lies not only in their capacity to describe patterns of actions and interactions of those studied but also in their potential to prompt one to ask questions regarding possible different perspectives and thus contribute to sharpening thinking that can equally contribute to professional development.

However, we are also clear on the necessity of understanding people’s contextual realities before introducing changes aimed at improving the quality of education in any context. To that, we add that when used in professional development programmes they can also suggest practices to the participants that are worth putting to the test of practice to

see if they lead to 'good' leadership on their part within their particular circumstances. We also hold that such suggestions from beginning school leaders are likely to be forthcoming when it is clear to them that what is being presented for cogitation is embedded in the realities of workplaces found in schools and in the environments in which they are located.

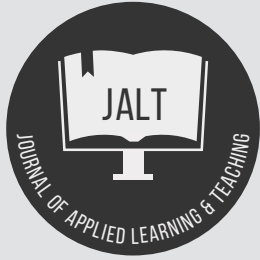
An example of such an invitation to readers is to explore the strong emphasis and value placed on the relational dimensions that emerge from the advice offered by principals in this small study. In turn, it might be of value to begin to explore the interplay between the well-being of principals and their ability to promote a culture of distributed leadership among their colleagues. A culture which in turn might instil both confidence and competence among the next generation of leaders and assist in making the position of principal more attractive than would currently seem to be the case. Yet, as always, such invitations are issued safe in the knowledge that one's current perceptions of one's context are key.

## References

- Blackledge, D., & Hunt, B. (1985). *Sociological interpretations of education*. Croom Helm.
- Bryant, A. (2013). The grounded theory method. In A. Trainor & E. Graue (Eds.), *Reviewing qualitative research in the social sciences* (pp. 120-136). Routledge.
- Clemans, A., Rickinson, M., Wilkinson, J., & Brooks, J. S. (2016). Principal preparation in Victoria, Australia: Conceptual foundations and design of the unlocking potential program. *Journal of Authentic Leadership in Education, 5*(1), 1-10.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research (3rd Ed.), Techniques and procedures for developing grounded theory*. Sage. <https://doi.org/10.4135/9781452230153>
- Da'as, R. A. (2023). Between principals' and a teacher's perspective taking: The role of transformational and transactional leadership styles. *International Journal of Leadership in Education, 26*(4), 722-744.
- Department of Education and Science, Ireland. (2022). *Looking at our schools*. Dublin: Stationery Office.
- Department of Education and Skills, Ireland. (2018). *Leadership and management in post-primary schools (Circular Letter 0003/2018)*. Dublin: Stationery Office.
- Education Act. (1998). *Government of Ireland*. Dublin: Stationery Office.
- Eisner, E. W. (1984). The art and craft of teaching. In J. Reinhart (Ed). *Perspectives on effective teaching and the cooperative classroom* (pp. 19-21). [https://archive.org/details/ERIC\\_ED250279](https://archive.org/details/ERIC_ED250279)
- Entwistle, H. (1971). The relationship between theory and practice. In Tibble, J. W. (Ed.). *An introduction to the study of education*. Routledge and Kegan Paul.
- Egley, R. (2003). Invitational leadership: Does it make a difference? *Journal of Invitational Theory and Practice, 9*, 57-70.
- Fullan, M. (1982). *Implementing educational change: Progress at last*. <https://eric.ed.gov/?id=ED221540>
- Hallinger, P., Gümüş, S., & Bellibaş, M. Ş. (2020). Are principals instructional leaders yet? A science map of the knowledge base on instructional leadership, 1940–2018. *Scientometrics, 122*(3), 1629-1650.
- Hargreaves, A. (1993). Collaboration: A key to leadership for quality in education. *Practising Administrator, 15*(3), 16-18. <https://search.informit.org/doi/10.3316/aeipt.79690>
- Henkel, T. G., Marion Jr, J. W., & Bourdeau, D. T. (2019). Project manager leadership behavior: Task-oriented versus relationship-oriented. *Journal of Leadership Education, 18*(2), 1-15.
- Kennedy, M. M. (1979). Generalizing from single case studies. *Evaluation Quarterly, 3*(4), 661-679. <https://doi.org/10.1177/0193841X7900300409>
- Jamshed, S. (2014). Qualitative research method-interviewing and observation. *Journal of Basic and Clinical Pharmacy, 5*(4), 87-88. <https://doi.org/10.4103/0976-0105.141942>
- Japhet, E. L., & Usman A. T. (2018). Factors that influence teachers' adoption and integration of ICT in teaching/learning process. *Educational Media International, 55*(1), 79-105. <https://doi.org/10.1080/09523987.2018.1439712>
- Lamm, K. W., Carter, H. S., & Lamm, A. J. (2016). A theory-based model of interpersonal leadership: An integration of the literature. *Journal of Leadership Education, 15*(4), 183-205.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.
- Morse, J. M. (2010). How different is qualitative health research from qualitative research? Do we have a sub-discipline? *Qualitative Health Research, 20*(11), 1459-1464. <https://doi.org/10.1177/1049732310379116>
- Murphy, G. (2023). Leadership preparation, career pathways and the policy context: Irish novice principals' perceptions of their experiences. *Educational Management, Administration & Leadership, 51*(1), 30-51. <https://doi.org/10.1177/1741143220968169>
- National Association of Principals and Deputy Principals. (2020). *Survey of senior leaders in schools*. <https://web.archive.org/web/20230521114342/http://www.napd.ie/less-than-one-in-three-secondary-school-leaders-predict-theyll-still-be-in-a-leadership-role-in-five-years-time/>
- O'Brien C. (2020, March 10). Warning over emergency shortage of secondary school principals. *The Irish Times*. <https://www.irishtimes.com/news/education/warning-over-emergency-shortage-of-secondary-school-principals-1.4197834>



- Organisation for Economic Co-operation and Development. (1991). *Reviews of national policies for education – higher education in Ireland*. <https://web-archiver.oecd.org/2012-06-14/72839-reviews-of-national-policies-for-education-higher-education-in-ireland.htm>
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd Ed.). Sage.
- Punch, K. F., & Oancea, A. (2014). *Introduction to research methods in education* (2nd Ed.). Sage.
- Purwanto, A. (2020). Did transformational, transactional leadership style and organizational learning influence innovation capabilities of school teachers during Covid-19 pandemic? *Systematic Reviews in Pharmacy*, 11(9), 299-311.
- Sachs, J., & Logan, L. (1990). Control or development? A study of inservice education. *Journal of Curriculum Studies*, 22(5), 473-481. <https://doi.org/10.1080/0022027900220505>
- Spillane, J. P. (2012). *Distributed leadership*. John Wiley & Sons.
- Stake, R. E. (1978). The case study method in social inquiry. *Educational Researcher*, 7(2), 5-8. <https://doi.org/10.3102/0013189X007002005>
- Stenhouse, L. (1975). *An introduction to curriculum research and development*. Heinemann Educational Publishers.
- Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Sage Publications, Inc.
- Uhrmacher, P. B. (1993). Coming to know the world through Waldorf education, *Journal of Curriculum and Supervision*, 9(1), 87–104. <https://eric.ed.gov/?id=EJ472515>
- Walsh, I., Holton, J. A., Bailyn, L., Fernandez, W., Levina, N., & Glaser, B. (2015). What grounded theory is... a critically reflective conversation among scholars. *Organizational Research Methods*, 18(4), 581-599. <https://doi.org/10.1177/1094428114565028>
- Weber, M. (1994). *Weber: Political writings*. Cambridge University Press.
- Wise, G., & Slater, C. L. (2020). *Critical moral leadership: Toward social justice for English learners* (pp. 51-68). Springer International Publishing.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin

---

Rose Luckin <sup>A</sup>	A	<i>Professor of Learner Centred Design, UCL Knowledge Lab, University College London, UK</i>
Jürgen Rudolph <sup>B</sup>	B	<i>Director of Research &amp; Learning Innovation, Kaplan Higher Education Academy</i>
Martin Grünert <sup>C</sup>	C	<i>Dean, Academic Services, Kaplan Higher Education Academy</i>
Shannon Tan <sup>D</sup>	D	<i>Research Executive, Kaplan Higher Education Academy</i>

---

### Keywords

---

Artificial intelligence (AI);  
AIEd;  
education;  
ethical AI;  
generative AI (GenAI);  
higher education;  
(human) intelligence;  
machine learning.

### Abstract

---

Professor Rose Luckin, a pioneer in the integration of artificial intelligence with education, holds the position of Professor of Learner Centred Design at the UCL Knowledge Lab, University College London. Her trailblazing research has profoundly deepened our understanding of AI in education (AIEd). Rose Luckin has authored over 50 peer-reviewed articles and key works, including "Machine learning and human intelligence: The future of education for the 21st century." As the Director of EDUCATE, she merges academic insights with ed-tech industry innovation. She is the co-founder of the Institute for Ethical AI in Education.

In our interview, Rose Luckin shares her educational awakening and her personal journey into AIEd, addressing gender bias and the unique challenges faced by women in the AI field. She delves into the ethical dimensions of AI deployment in educational settings, underscoring the Institute for Ethical AI in Education's pivotal role in fostering ethical standards. Professor Luckin advocates for AI's potential to bolster learner-centred methodologies and stresses the critical importance of forging robust partnerships between educators and technology developers. She evaluates the impact of generative AI on assessment, learning and teaching within K-12 and higher education. She provides insights into AI's evolving role in education and the imperative of lifelong learning. Emphasising a collaborative ethos among educators, researchers, and developers, Professor Luckin argues for AI's integration into education within strategically crafted ethics and governance frameworks. Our interview sheds light on AIEd's current landscape, highlighting the critical need for ongoing research and collaborative efforts in navigating its considerable dangers while seizing opportunities.

---

### Correspondence

Jurgen.Rudolph@kaplan.com <sup>B</sup>

---

### Article Info

Received 6 March 2023  
Received in revised form 18 March 2024  
Accepted 18 March 2024  
Available online 26 March 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.27>

## Personal experiences of school, further and higher education

**Jürgen Rudolph (JR):** You are such a distinguished figure in the burgeoning field of artificial intelligence and education. You are the Professor of Learner Centred Design at the UCL Knowledge Lab, part of University College London, where you have been an influential force for over two decades. Your pioneering research significantly advances our understanding of AIED.

You are not only a prolific academic, authoring over 50 peer-reviewed articles and several notable books, including "Machine learning and human intelligence: The future of education in the 21st century" (Luckin, 2018), but you are also a leader in applying research practically. As the Director of EDUCATE, a hub for educational technology start-ups in London, you have fostered innovation and collaboration between academia and the ed-tech industry. Your expertise has been recognised widely, including your appointment as a Specialist Adviser to the UK House of Commons Education Select Committee and your role as a Co-founder of the Institute for Ethical AI in Education. Before joining UCL, you served at the University of Sussex in various capacities, including as Pro-Vice-Chancellor (Teaching and Learning) and Director of the Human Centred Technology Research Group. Your administrative and research roles have been instrumental in shaping policies and strategies in both teaching and technological contexts.

To the best of our knowledge, academia was your second career. Prior to the birth of your first child in 1983, you had a career in Banking and Finance. You had been a Top 20 candidate in the UK for the A.C.I.B. Chartered Institute of Bankers Associateship Examinations. In 1990, you started to pursue a BA (Hons.) in Computing and Artificial Intelligence at the School of Cognitive and Computing Sciences at the University of Sussex (you were awarded a First Class Honours), and between 1994-1997, you pursued a PhD with ECOLAB on Explorations in the Zone of Proximal Development at the School of Cognitive and Computing Sciences, University of Sussex from 1994 - 1997. Could you please tell us more about your own schooling and university education?



Figure 1: Rose Luckin at a conference in China in 2019.

**Rose Luckin (RL):** It's a very interesting question on many levels because my own schooling was quite disrupted. I didn't have a smooth journey through the education system. I had a lot of problems when I was in my early teens. I was a school refuser, though I was very good at sports. It was a very insightful sports teacher who sat me down one day and said, 'You know, you could be good at something else as well as running'. It's funny how one person can really have quite an impact because, in the UK, we have a system that involves selection at 11, called the eleven-plus [a competitive examination given between primary and secondary school at about age 11 as a means of determining in which of the three types of secondary school – grammar, technical, or modern – a child should continue their education]. I had failed the eleven-plus when I was expected to pass it.

My parents were very dismayed. I've let them down, and part of the reason that I was struggling at school was because I felt like a real failure. Obviously, my parents had confidence in me, but I felt too guilty for letting them down. So, it was very important when the sports teacher showed confidence in me. I've been very frank here, but I think it's quite interesting for your readers to realise that you can become a professor even when you've had quite a disrupted early education. Also, this disruption has impacted a lot on my research interests.

*You can become a professor even when you've had quite a disrupted early education.*

So, when I was encouraged to take learning seriously, I was very fortunate that my parents moved me to a different school to give me a fresh start. It was a shame because it meant I left the PE [physical education] teacher, but it did give me a fresh start. I then, of course, worked super hard. It was like going from one extreme to the other to try and catch up. In the end, I did get my formal qualifications, my O-Levels – now GCSE [General Certificate of Secondary Education] – and the right kind of grades. I did get my A-Levels, and interestingly, I ended up in the same school I would have gone to had I passed the eleven-plus to do my A-Levels. So, it came full circle in the end.

But even then, I didn't go to university straight away because I came from a family that had never had anybody go to university. Despite that fact, my father was a lovely man, and I loved him dearly; he's sadly dead now. He basically said to me that when I said 'I'd like to go to university': 'No, you need to get a job; that's what we do. You've done well, really well. But let's get a job now, my love.' This is the same man who, despite having no expertise in the area of my PhD, read every word of my PhD to check the punctuation [all laugh], so he is a lovely man.

But that was the way it was. So, I didn't go to university when I left school, despite having the A levels that would have enabled me to do that. So, I went to the bank because I needed to get a job. I was the only woman taking the exam in the regional branch where I was working. They were a bit surprised when I said I wanted to do the A.C.I.B. Chartered Institute of Bankers Associateship Examinations qualification. I didn't just want a job; I wanted a career. But they agreed.

Bless them! They supported me through college because I had to do it as a day-release student in a further education college. You would work, and then, every other week, you would have a day where you would go to college. But it was hard because you were doing a lot of studying at the same time as holding down a full-time job. But that was good for me.

I quite thrived on that concentrated effort. I surprised them when I did really well in my exams, finished in the top 20, and got a little medal. The branch was excited. They've never had anybody do this, so they were very pleased. It was a very interesting educational experience for me in the further education sector and definitely formative to me as an educator. The secondary school education I had was initially in a very poor-quality secondary school. So, when I had completed the banking exams, I stopped full-time work and had my children. At that time, my husband was working as a schoolteacher, so we didn't have a huge amount of money, but it was fine. I took full maternity leave, and I didn't go back to work other than part-time until both my children were at school.



Figure 2: Rose Luckin and daughter.

I became a tutor for something initially called the Rapid Results College, which was a correspondence college. Again, this was very influential on the research I later conducted. Before the Internet, I would receive students' work in the

post, mark it, and send it back to them with feedback. Then, they promoted me so that I was writing the course textbooks. The students who enrolled in Rapids Results College got a set of texts to help them through the banking exam content. I was writing the little course textbooks, pointing them to other resources and structuring the instruction in particular subject areas. That, again, very much informed who I then became.

I then – in order to bump up the family budget, to be honest, but also to keep my brain cells working – became a part-time tutor at the local further education college. So, after having benefitted from the further education system in the UK, I then became a teacher in the further education system. That was really interesting because I was engaged in teaching students, not just the banking subjects, which is obviously what I had been tutoring for Rapid Results College.

But I was also asked to teach adults with special educational needs. I have no training in doing this at all. This was incredibly informative to who I then became as an educator: to be set the challenge of 12 very significantly disadvantaged young people, trying to help them use a computer. In one instance, the student didn't even recognise that when he pressed the letter on the keyboard, he got the same letter on the screen. We're talking about serious disabilities in learning terms, so it was very challenging. At that time, I realised that because the legislation around banking changed in the UK, I wasn't going to go back to a career in banking because it wasn't the same job that I had left. A lot of the more interesting activities in local banking had been moved to central offices and taken out of the regional branches, and I didn't want to end up in some kind of service centre. So I thought, 'Well, ok, I don't have a degree, let's now go to university, let's have a look'.

Because I had small children, geographically, there was only one university I could get to. It was quite easy to fill out the form, and I had the A levels and the banking qualification. I thought I'd probably better do something financial or economic. I applied to do economics. Then, I looked through the student brochure and saw this subject called artificial intelligence and computer science. I had never heard about it, read through it and thought, 'Oh, this looks interesting'. They recommended some reading, so I went and bought the books. One of which was "Gödel, Escher, Bach" [Hofstadter, 1999; first published in 1979], which is an incredibly difficult read. It was a very intense introduction to AI philosophy, things I hadn't engaged with previously but fascinating, and I loved it. I thought: 'Nothing to lose. I'll just put it down on the application form and see what happens.'

I got rejected from economics and accepted for artificial intelligence and computer science. When I became Pro Vice Chancellor, the person who was the vice chancellor was the man who rejected me from economics, and I never let him forget it. [All laugh.] He rejected me because he thought I had already done it. He thought my banking qualification would have influenced my thinking. So he said, therefore, I wouldn't approach the subject fresh. But he did me a favour because I loved computer science and AI.



Figure 3: Professor Rose Luckin.

I was one of two female students on my course. Everybody else was male. I went to Sussex University, which encouraged mature students. There were other people who were more of my age because I was in my 30s at this point. But they were male, and at one point, the other female dropped out of the course. So, I was the only female in the cohort. I had a very different experience of the university and degree study because I couldn't join the social things, as I had two small children.

I had to get home, and I had a very intense work phase. I would work in the middle of the night because it was the only time I could get quiet when the children were in bed. We lived in a very rural village with no street lighting. So, I would often sit in the bedroom working away and look out. It'd be completely black outside. It felt quite alone trying to work out this computer programme that wasn't working. A lot of the time was spent debugging.

Anyway, that's how I got into computer science and AI, and it was interesting being female. But, in the end, I had a wonderful relationship with the other students, even though most of them were younger than me. Once they realised that I could actually do the work and that I was a serious learner, we got on great. I'd help them. They'd help me, too. It was a really lovely learning experience, and it was a very interdisciplinary course that has very much influenced the way I think, too, because it was quite the early days of AI. This was 1990, so it was the very early days of AI, and

it was very much about what intelligence is. So, we were looking at psychology and philosophy, and of course, we did theoretical computer science and programming and all of those things. But we also did linguistics and lots of contextual courses, and I really enjoyed it. I loved studying, and I loved the course. That was a very long answer, but I hope it gives you some useful context that your readers might find interesting when thinking about how you get into a subject like AI.

## Gender bias in academia

**Shannon Tan (ST):** You have researched AI since the 1990s. As one of the early female pioneers in the field of AI and education, could you share your experiences regarding the challenges you faced and overcame? We suspect you encountered glass ceilings that you have successfully shattered during your stellar academic career. How have you seen the gender dynamics in this field evolve over the years?

**RL:** It's very complex. I definitely experienced challenges both as a female and as an older female. Because of my previous experience, I initially regarded some of those challenges as being my own fault. So, for example, I would never use my children as an excuse for not being able to do something because I didn't want to draw attention to the fact that I was female and had children. Of course, everybody knew I was female, but I didn't want to use that as an excuse for not being able to meet a deadline or not being able to do something. I would make a huge effort not to have to do that. I don't know whether some of the challenges I faced, particularly in the early part of my career, were to do with me being female or to do with me being quite successful quite quickly. It's hard to know why people behave badly towards you because there's a lot of competition in academia. There are always people who want to get one over on you. I wouldn't want to blame gender for things that might not be gender-related.

But what I would say is that things have improved in terms of gender politics and the glass ceiling. It's still an issue and a challenge. I don't think we've solved it. In fact, in later years, I would say that the challenges that I experienced came more from other females than from males. Unfortunately, perhaps because it is difficult to be a successful female academic, women don't always support each other as much as they could. It's a very complex issue, and there is no simple answer. But I certainly have experienced challenges. The way I tended to try and deal with them was to deny that they had anything to do with the fact that I was female. Later, my approach has been to try to be as collaborative as possible, to try and accept that you can't agree with everybody and that there will be people who will do you down. It's a fact, but you try not to take it too personally, see it as the pressure they're under, and emphasise collaboration, which is why so many of my publications have many names on them. This is not always well respected, but it very much speaks to that collaborative approach that I've found personally the best way of navigating what is a tricky situation.

**JR:** I agree that collaboration is brilliant. Unfortunately, when you look at the metrification at universities, they oftentimes very much favour single-authored publications.

## Ethics in AIEd

**Martin Grünert (MG):** As a co-founder of the Institute for Ethical AI in Education, you have been at the forefront of discussing ethical considerations in AIEd. What do you consider the most pressing ethical challenges in implementing AI in education, and how can educators and technologists work together to navigate these challenges?

**RL:** We have to work together because it's so complicated. When I formed the Institute with [Sir] Anthony [Seldon] and Priya [Lakhani, OBE], the bottom line of why we formed it was because we were really worried that nobody was paying attention to education when it came to AI. There was lots of data and AI, lots of ethics and AI, and ethics and data. But nobody was looking at education. I believe to this day that education is a special case because we want people to engage with education throughout as much of their lives as possible – unlike the medical system, where we want them to engage with it as little as possible because they're healthy. It also involves vulnerable people. The reason we founded the Institute for Ethical AI in Education was that we were concerned that nobody was really paying attention and that if we didn't do something, all of the great benefits that AI could bring to learners, teachers and parents would be lost because there'd be some ethical dilemma or horrible occurrence and everybody would, of course, have a knee jerk reaction against the technology.

A large part of the motivation for founding the Institute was to draw attention to education as a special case and to try and alert people to the fact that we needed to think very carefully about ethics. Of course, the situation's got even more complicated with the launch of generative AI applications that have made open-source, scaled, and sophisticated AI available to anybody. That means that people with bad – or good – motives can develop sophisticated AI systems, and that's dangerous. So, we have to work collaboratively and engage all stakeholders around ethics. It's not just about the technology itself; it's also about how the people developing the technology can engage in ethical codes of practice. It's not just about the regulation; it's also about the education of the general public. From my perspective, a key section of that is that educators need to understand more about AI, what it can do, and the basics of how it works. They don't need to know the deep tech complex science because not everybody's into that, but they do need to know the basics.

We have a problem because (a) the regulation and the code of practice will never keep up with what the technology is able to do, and (b) huge assumptions will be made about, for example, what it means to be transparent. Because transparency is not a one-way process. You might think that as a developer, you have explained what happens with your data very carefully. For example, if you're using machine learning, what data do you collect, how do you process them, and how is your system trained all of that time? Unless the person reading it has a certain level of understanding,

it's not transparent to them. So, there's a huge educational imperative around that ethical space. It absolutely has to be a collaboration between regulators, developers, educators, and community groups. It's really important that we see the immense impact that these technologies have on society and, in particular, on education. We're still waiting to hear when GPT-5 will be released, and we're being told lots of things about what it's going to be able to do, and quite frankly, I find it frightening what I'm hearing (see, e.g. Shah, 2024). Is the world ready for this? It's quite a worrying space.

## Advocating educator-developer partnerships

**MG:** Indeed, there's always a risk if the educators are not involved in the process. As we are discussing education, it's concerning to me, as an educator, that solutions are often designed based on perceived needs rather than what will truly be effective and positively impactful. The next question is about the symbiosis between educators and developers and how they can work together. How do you see that partnership being able to be firmed up or strengthened in order to be able to generate the most beneficial outcomes?

**RL:** The bottom line is that it takes time, and it takes patience and a willingness for all parties to perhaps put aside some of the things that they might normally prioritise. Let me be a bit more specific. I've written in the past about the importance of AI developers working with educators (e.g. Luckin et al., 2016; Cukurova et al., 2019; Weatherby et al., 2022) – and I do believe it absolutely is the way forward. But it's a very difficult road. When I was talking about my history when I was teaching at the further education college, I didn't mention that I also taught in school for a short period of time, again in a very challenging environment. I don't think I was very good, but it has given me an understanding of what it's like to be a secondary school teacher with a very challenging group of young people in their teens.

When you've got multiple things going on, and you've got to keep control of and support that group of people, the last thing you want is a not-quite-finished bit of software to work with. You need something you can depend on. You have to do it in stages. There is a lovely book chapter called Lesson 21 by Kate Erricker, with whom I worked on a project for Nord Anglia Education, that really highlights this so nicely and tells the story of trying to co-design a piece of software and how incredibly difficult it was for the teachers, even with the best will in the world, to do this in a high-stake learning environment (see Erricker, 2023). In those learning environments, it's high stakes in a different way compared with my situation with very challenging learners.

You can't work with something that's still being built, and yet the only way to really engage educators in that process is for them to be part of that design. You have to have stages. Engaging educators in being able to be part of pilots, where they're not necessarily operating in that high stakes environment whilst the technology is not in a place that would support them in that environment but helping them to engage in the design process to understand a bit more about what it is, and how AI works.

If you can help people to understand a bit more about what it is that AI is doing and how AI uses data and what that means, they then are much more able to become proficient users and also understand what potential ethical challenges and safeguarding challenges might be. So, it needs to be long-term relationships. They need to be patient. People need to be absolutely upfront about the challenges of the activity right at the outset. Initially, it needs to be done in a way where teachers aren't put in high-stakes environments with semi-functional technology because it's just too stressful. But they can be part of projects, and then once you get to a certain stage where they feel comfortable and the technology is in a state, then you can start rolling it out and testing it in those more real-world, authentic, high-stakes environments.

### AI's potential in enhancing learner-centred approaches

**JR:** Could you please walk us through some of your key works? In your book "Machine learning and human intelligence: The future of education in the 21st century" (Luckin, 2018), you discuss the potential of AI to transform education. Could you elaborate on how AIED frameworks can be optimised to enhance both teaching and learning experiences, especially in the context of fostering critical thinking and creativity in students? This is related to your 2010 book "Re-designing learning contexts", where you advocated for learner-centred design in education. How do you see AI further enhancing learner-centred approaches, and what gaps in research and practice still need to be addressed to achieve this?

**RL:** The whole learner-centred design piece has been a core part of my thinking throughout all of that work. When I was writing the book on "Re-designing learning contexts" (Luckin, 2010), I used writing as a way to help myself understand more and help myself learn, which is why I find the suggestion that an AI would ever take over my writing process completely meaningless. Because I actually want to learn from trying to express myself. I'm sure AI can help. Don't get me wrong. But we learn from exposing our understanding, examining it, reframing it, learning from it, and challenging it. I think the framework that I put forward in the "Re-designing learning context" book is something that I'm now revisiting because AI itself is now far more able to tap into those contextual features that I was trying to identify in that book. So, the primary reason I wanted to write that book was to try and understand more about it - How could we talk about learning contexts?

There's a huge literature on context, but it's quite complex. There are lots of different disciplines that talk about context, and they talk about it in quite different ways. When I was writing that book, I concluded that, from my perspective, being pragmatic and trying to come up with a framework could be useful when thinking about designing the way that learners interact with technology. It was more useful than conceptualising context as something that people step into or out of: 'I'm in this context now' or 'I'm in a learning management system' or a 'virtual learning environment context for a moment'. 'I'm in my office context'. Another moment, 'I'm in the car context'.

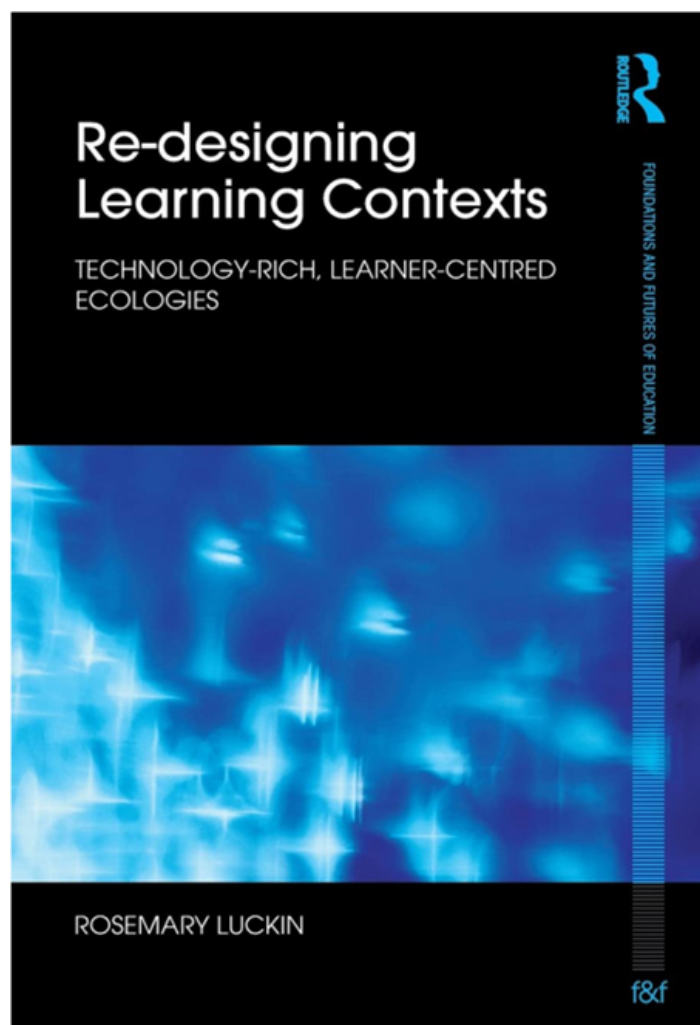


Figure 4: Re-designing learning contexts (Luckin, 2010).

I saw those as environments, and I saw context itself as a whole set of interactions that I might have with the world as I go through my life, but with a particular slice of learning. I believe that people create their own context from all the interactions they have in the world. But of course, that's huge. How do you start thinking about that from a design perspective? Then I tried to think about how you might think about ways in which you could identify different elements of that context. This is an oversimplification. So, thinking about digital and physical environments – what are they like? I thought about the kinds of tools, both technical and non-technical, that you use within their environment and how much there is a relationship between the environment and the tools that are used.

For example, in a science lab, you use quite different tools from the tools that you use in an English language classroom. This is quite an obvious connection, but there's a connection between the nature of the tools that are used and the environment, and then to think about what it is that you're trying to learn. So, you have your learners in the centre, and they interact in different environments. They use different tools, and they access knowledge, information, understanding, skills, and capabilities about a particular subject area or skill through their interactions in environments. But none of those things are necessarily a direct relationship.

So, a child in a classroom doesn't necessarily have the ability to interact with the teacher whenever they want. There are written and unwritten rules about how that works. So that's why, in the framework of that book, there are these items called *filters*. They can be positive or negative in terms of the way that they constrain the access that a learner has to a particular resource. That resource might be part of the environment, a tool, or a part of what they're trying to learn. So that's where that book came from. It was a way of trying to take a very complex subject like context and turn it into some kind of framework that could be useful when we're thinking about designing learning experiences.

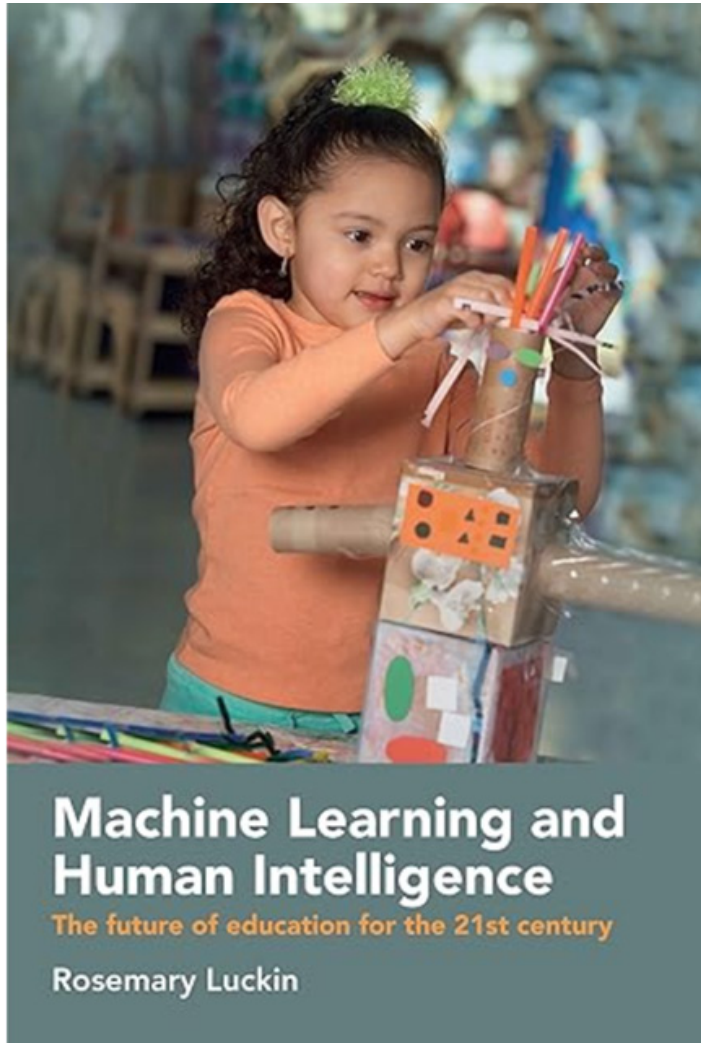


Figure 5: Machine learning and human intelligence (Luckin, 2018).

Now, with the ability that AI has to capture data about those interactions that a learner is having as they progress through their journey, I'm relooking at that work: it's interesting to think about how we can consider design frameworks for AI that are informed by that conceptualisation of context. Then, the other book, "Machine learning and human intelligence" (Luckin, 2018), was written explicitly because I wanted to understand how better I could talk about human intelligence in a way that would distinguish it from artificial intelligence. Because even back in 2017/18, when I was writing that book, I was really bothered about the way we were talking about AI as if it was human-intelligent when it wasn't, and it still isn't human-intelligent. Not the same way that we are.

*I was really bothered about the way we were talking about AI as if it was human-intelligent when it wasn't, and it still isn't human-intelligent.*

So, that was the absolute motivation for writing that book. That's why the framework in that book is very much focused on how we can differentiate AI from human intelligence and looking at what's happening with the large language models, for example. Particularly now, they're not just language; they're multimodal. I'm having to revisit that and challenge myself about how those differences are perhaps being blurred. Whilst I don't believe that the sorts of generative AI models we're seeing understand the world in the way that we do (because they don't directly experience the world), I do see that there could be an argument that says: 'I spend six months interacting with a multimodal generative AI model. In those conversations, I can see that there's a reasonable premise that you could consider that that model is building an understanding, a representation of the world that we are interacting about'. It doesn't have direct experience of the things that we might discuss but I think there is some notion. There are still clear differences between human and artificial intelligence, but the boundaries are more blurred.

*There are still clear differences between human and artificial intelligence, but the boundaries are more blurred.*

It shows that whatever you write, you have to revisit it all the time because our understanding of human intelligence, our understanding of AI, is changing so fast. So, if you write something one day, you have to be willing to revisit things. So, I still believe in the fundamental premises of both books, but I think the precise detail is shifting, and that's really interesting.

**JR:** Everything is developing at breakneck speed. This can be exemplified by Sora, OpenAI's text-to-video AI (see Brooks et al., 2024) that created global headlines and will soon be launched. It's staggering, and it raises questions about what truth is and what reality is. As you were saying earlier, it's quite scary. By the way, we are currently working on an AI framework for our institution's teaching and learning, and we are quite inspired by your ethical AI framework (Luckin et al., 2022).

**RL:** It's really useful to know that because we went to a lot of trouble to try and make the output from that work really practical.

### The Golden Triangle Framework

**JR:** In this context, your Golden Triangle Framework (e.g. Luckin & Cukurova, 2023) provides a model for AI and education by connecting three key elements – educators, researchers and EdTech developers. The Golden Triangle Framework emphasises the role of the teacher in orchestrating the educational experience, selecting appropriate AI tools, and ensuring they are used effectively. AI can provide adaptive learning experiences personalised to each student's needs,



## The 7 steps to AI Readiness: ETHICAL

There are seven key steps to getting your organization ready to leverage the transformational power of AI. These can be found in the 'ETHICAL AI Readiness' framework:

- 1) **Educate, enthuse, excite** – about building within your community an AI mindset
- 2) **Tailor and hone** - the particular challenges you want to focus on
- 3) **Identify** – identify (wisely), collate and
- 4) **Collect** – new data relevant to your focus
- 5) **Apply** - AI techniques to the relevant data you have brought together
- 6) **Learn** – understand what the data is telling you about your focus and return to step 1 until you are AI ready
- 7) **iterate**

And all these steps should be done ethically



Figure 6: The seven steps to AI Readiness (Luckin, 2020).

giving real-time feedback and enabling practice through repetition. However, good AI education solutions require human-centred design that is focused on supplementing/enhancing the teacher's role rather than replacing them. Effective AI requires an iterative, evidence-based approach drawing on multidisciplinary expertise in learning science, technology, and practice. AI should be designed ethically with transparency, privacy and agency in mind to foster trust and metacognition. Your framework highlights the need for research on how best to integrate AI in education to improve learning outcomes. Would it be fair to say that AI should not be on autopilot but, at best, a co-pilot? Could you also please elaborate further on your Golden Triangle Framework, perhaps sharing aspects or nuances that are sometimes overlooked?

### The Golden Triangle

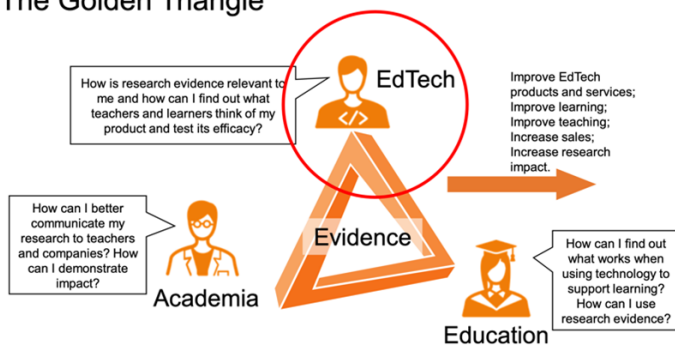


Figure 7: The Golden Triangle Framework (Luckin, 2020).

**RL:** That's a very big question. I will start with the Golden Triangle, but I really want to come back to that question about co-pilot versus autopilot because I think it's the crux of where we are with AI. Let me give you a sense of the context of the Golden Triangle. For many years, when I've been working in what is essentially AI and education, the word 'AI' has not been popular. 15 years ago, you didn't necessarily want to say you were working in AI because people didn't necessarily want to hear it. So, we would talk about educational technology, and of course, some educational technology doesn't use any AI anyway.

*For many years, when I've been working in what is essentially AI and education, the word 'AI' has not been popular.*

Although I've always been doing research on AI in education, I did branch out to look at educational technology more generally because of my own personal learning experiences and my focus on applied research. I was always looking to do things that would actually have a useful, practical application. At that time, most of the things that were being used in education were educational technology, not AI. I spent quite a bit of time thinking about and looking at different educational technologies and talking to companies who were developing them. I've always believed that having a relationship with those who are developing their technologies commercially is really important. About 12 years ago, on January 6, I pulled together a meeting and had a wonderful facilitator. It was a meeting where I brought together some big tech companies, some small tech companies, and some educators from across the different sectors. Basically, the question I asked them to answer is: 'Why are we not better able to use technology in education impactfully? What is the problem here?' We came to the conclusion that there was just a really big disconnect.

*'Why are we not better able to use technology in education impactfully? What is the problem here?' We came to the conclusion that there was just a really big disconnect.*

You had a whole body of people who were researching technologies for use in education and sometimes building research tools that never made it to commercial viability. Another bunch of researchers were exploring the use of existing commercial tools in education but very rarely actually speaking to the people who built the technologies. Then you have this whole raft of people who were building technologies, both small-scale and large-scale, who didn't know how to contact the right person in the research world to create the right connection. Then you had the poor old educators who were left at the mercy of both groups, quite frankly.

We concluded that if we could do one useful thing, it would be to better connect all of those stakeholders. My passion for collaboration is not surprising. The group also thought that we needed to think about how we could create better connections between the communities of people who built the technology and use it. This could help us understand how, why, when, and for whom it is usefully impactful. That's where the Golden Triangle came from.

Then, unfortunately, I was diagnosed with breast cancer, and had a rough ride. If I'm honest, because I had to have all the treatments, the chemotherapy, the surgery, the hormone therapy, everything, I was very ill for two years. I completely lost two years.

**JR:** I'm very sorry to hear that.

**RL:** You probably won't want to put this in the article, but I don't actually mind if you do if you think it would be of interest. I also had a complete psychological breakdown and was in a secure psychiatric institution for a month. The reason I am willing to share that is because it has made me think

about artificial intelligence differently. Because my mind was completely and utterly broken in a way that is really hard to explain. And yet I recovered. That's amazing. It's an amazing capacity of the human brain and our intelligence to recover from such a fundamental breakdown where you were such a danger to yourself that you have to be locked up, literally.

*My mind was completely and utterly broken, and yet I recovered. It's an amazing capacity of the human brain and our intelligence to recover from such a fundamental breakdown.*

The reason I'm always willing to share it is because I am a very honest person, but also because it really informed the way I think about artificial intelligence and human intelligence. There's so much we don't understand about our human intelligence and capacities that I hope we will learn more about because we have these AI tools. I hope we'll see it as an opportunity to really explore the depths of our capacities. That was a hard time. I lost two years completely and probably, if I'm honest, three because it took me a long time to get back physically and psychologically to be able to cope with a full-time and quite high-pressure job. But I did get back at full throttle, as you can see. That's why I say it's an amazing capacity of the human mind.

But it meant that that conversation just got left there. Then, three years later, I thought about, 'How do I pick this up?' I started to reengage with the people, and interestingly, they were all happy to reengage in the conversation. Nothing had happened to them in the meantime, really. There had been a small group who tried to take some of the ideas forward but hadn't really kept its momentum. We revisited it, and then, to cut a long story short, at the end of 2015 and the beginning of 2016, I was encouraged to apply for some funding. It wasn't research funding as such. Development funding from the European Union was used to build better relationships between small companies' entrepreneurs and, in my instance, Edtech companies and universities. So, I put in a proposal for this funding to try and explore this Golden Triangle concept that we'd come up with, and we'd come up with a name – EDUCATE.

Fortunately, that was funded. But of course, a few months after I got the information that it had been funded, the UK decided to leave the EU. That was another little barrier in my way [all laugh]. We have to wait for various political ramifications. Still, in the end, the funding came through and in 2017, the EDUCATE programme started. That had the explicit goal of trying to bring together those three communities. It wasn't about AI. It was about educational technology in general. Of course, AI is relevant and fits within that, but it would be part of the educational technology category.

We did have companies like Century Tech, which is quite a big AI company, that came on one of our very early cohorts and what we did was quite simply design a training programme to try and help developers better understand how to evidence what was impactful about their particular product or service, and to start building a logic model.



Figure 8: Rose Luckin at UCL EDUCATE Demo Day 2019. Source: Team EdTechX, 2019.

They were clear about what the educational outcome they wanted to achieve was, what the outputs from their particular product or service were, how they could collect data and evidence and how they could analyse it along the way to try and help them have more informed conversations with their customers, the teachers. We wanted to engage researchers in this, and we did. Therefore, we also trained some researchers to understand better how to engage with these small tech companies because they both work to very different rhythms of life. You actually have to build that relationship as it doesn't happen naturally.

We've never managed to do the same relationship building at scale because we work with over 360 companies, which is a lot of the UK tech sector. Some companies were from outside the UK, but most are UK-based. We never managed to build the same depth of relationship with educators because it was so hard to access them. Because they're incredibly busy, don't have time, and can't use semi-functional products. Then, there's a lot of complexity for educators. We did engage with educators, but it felt that it was more from the perspective of giving them information than actually helping them to be part of a collaboration. It felt much more that we were exposing them to information and opportunities, but I never felt we'd built the kind of in-depth collaborations I would have liked to have seen. I'm still trying, obviously. Now, the Golden Triangle is highly relevant to the AI space. The building of the relationship between the different parties is fundamental to the success of any AI use case that's being rolled out.

In Singapore, there are three use cases being rolled out across schools. That was decided and designed before ChatGPT was launched back in 2022. It's really interesting to see how that unfolds. But we certainly need to build those relationships that are defined by that Golden Triangle and try to learn from some of the approaches to training, for example, those that were developed as part of that programme. That's certainly what we're trying to do at the moment. The company Educate Ventures' researchers are absolutely helping educators think about the purpose for which they want AI to be used. What do they want to

achieve with it? Rather than starting with the technology, asking What's an educational challenge? What is the task that could be helped by AI? To start from that point and then develop small tests, even before pre-pilot, even at a very early investigation of what it means for that AI to be used: What kind of infrastructure do you need to have? What kind of staff capability do you need to have? All of those are very practical things that you need to understand before you can get into really interesting questions about how that aspect of AI is supporting learning or how it is helping teacher development.

You need to get those pragmatics right. We have a little framework for doing that and a framework for thinking about how you evaluate each of the ways in which you are looking at AI and how you can take a very early intervention through to a pilot and think about how that can inform a bigger strategy. What I'm seeing, certainly in the UK but not exclusively in the UK across the entire education sector, is that people, including educators and students, are using AI. But I've not yet seen anyone who's developing it in an organisational, strategic way. What I've learned from the Golden Triangle is that we can build those relationships within an underpinning governance and ethics framework, which is fundamental in running a series of use cases of AI.

We look at the technical and data infrastructure required, how they build staff capability, and what staff capability is needed. We have a framework for learning from every iteration that feeds back into the strategic way in which AI is being used in an institution. That's the sort of framework we work with, and it's all really been inspired by that Golden Triangle. It's interesting how it's evolved, right back from 12 years ago.

**JR:** Thank you so much for sharing, and we're so sorry to hear you went through such a torrid time.

**RL:** I'm fine now. I'm always willing to be honest about it because (a) I think it helps other people who might have had similar experiences, and (b) it really taught me something about intelligence and the human mind; it really fundamentally shifted my own perceptions.

**JR:** As you know, Shannon and I co-authored a book with Stephen Brookfield (Brookfield et al., 2023). Stephen went through some ten years of severe clinical depression. For the same reason that you mentioned just now, he also talks about it very openly because he believes it's a taboo topic that men especially like to suppress because men always see the need to perceive themselves as strong and so on (Brookfield et al., 2023). I think it is excellent that educational thought leaders like you and Stephen Brookfield talk about their personal crises and how they eventually overcame them.

**RL:** I think it's really important to be open. I understand the male perspective on that, and I don't deny anything that he's saying, obviously. But that was part of the challenge that I faced as a female: not wanting to show any weakness. I tried to work throughout all my treatment, which was a complete mistake. The day before I collapsed and was admitted to the psychiatric hospital, I was trying to read a student's thesis.

I couldn't process the words, but I just didn't feel I could give in. I thought I had to keep the flag flying. We females might get breast cancer, but we can hack it! And actually, I couldn't. That's the honest truth.

*That was part of the challenge that I faced as a female: not wanting to show any weakness. I tried to work throughout all my treatment, which was a complete mistake. The day before I collapsed and was admitted to the psychiatric hospital, I was trying to read a student's thesis. I couldn't process the words, but I just didn't feel I could give in.*

## Generative AI and its impact on education

**ST:** Since its inception more than a year ago, ChatGPT has fired the public imagination with a vengeance. Within a couple of months, the AI chatbot has hit more than 100 million users. There is an increasing number of rival chatbots. Having been so long in AIED, you may be amused by the sudden hype that surrounds large language models (LLMs) and generative AI. What are your personal experiences and impressions of these developments in the past year? How do you see these advancements influencing future research and education in AI?

**RL:** I love the way you phrased that: "fired the public imagination with a vengeance"! On the one hand, I find myself completely delighted because people actually want to talk about AI. I've been trying to get people to want to have conversations. I've been trying, particularly, to talk to teachers, people who are actually practitioners. I'd been wanting to have much more conversation with them and school leaders about AI, and it was really hard to get anybody interested. Now, it's not hard at all, so it's good from that perspective.

It's a very powerful technology. I find it fascinating, and if I can manage to be detached, I can look at it in a more scientific way and think about: 'Okay, so what's the sort of sociotechnical story here?' We've got this technology. People weren't ready for it. It was a quite deliberate launch of something that society wasn't ready for. As a mass experiment on millions of people, that is interesting, but it's also quite frightening. I find myself on the boundary, often between excitement and fear, because I think it's great that people now want to engage with AI. Something like this was always going to happen. There was going to be a moment when the power of AI became more apparent.

*It was a quite deliberate launch of something that society wasn't ready for. As a mass experiment on millions of people, that is interesting, but it's also quite frightening.*

In a sense, the power of AI had been apparent but hidden before. Social media is driven by AI algorithms, and we know how much of an impact that has had on society, education,

and people. But it was not explicitly, 'I am a user using AI', 'I'm using social media, and I may or may not have any idea that there's AI actually running this'. So, this was an availability at scale of a very powerful AI that was easy to use, and you knew you were using AI. It had it labelled on the tin – 'OpenAI'. This is an AI. It wasn't hidden at all, which is all good. But of course, the world is not prepared for it, and that's my biggest concern. My experience to date tells me that I'm right to be worried, from politicians to schoolteachers to all sorts of different ages. The vast majority of them have no idea what they're doing with AI.

*In a sense, the power of AI had been apparent but hidden before. Social media is driven by AI algorithms, and we know how much of an impact that has had on society, education, and people. But it was not explicitly, 'I am a user using AI'.*

Understandably, why would they? It's not a criticism in the least. But I often get asked by policymakers, and they're doing their best. Some of the questions they ask are incredibly naive and really worrying because that's the kind of thing that is driving the way education systems work. Singapore is a different case because they've been engaging in thinking about AI much longer at a system level. I think they are definitely ahead of the game, but nevertheless, it's still a challenge to think, 'How do I help so many people grasp enough about AI to understand how and what they should use it for?'

This comes back to the question about co-pilot and autopilot. There are many reasons why I think that everybody needs to understand enough about AI to use it safely and to their benefit. They need to understand enough to build AI. Few people do, of course, but we don't need everybody to do that. It's because we don't have a good conceptualisation of what the relationship between an AI and human intelligence should be. There is no clearly defined framework for thinking about that relationship. There are no recommendations. We certainly need to start conceptualising the different sorts of relationships that can exist between artificial and human intelligence quite quickly and thinking about the ramifications of those. Because we already know that technology is changing the way our brains work at the neural level. We are changing because of the way we use GPS, because of the way we use Google rather than remember things.

*Technology is changing the way our brains work at the neural level.*

So, however we use this AI, it will change us cognitively. We better make sure those changes are the ones we want and that we don't offload the wrong things to the AI. That autopilot/co-pilot thing is fundamental. Unless we have an educated enough population, how do they navigate that very challenging space? Even, say, if it was possible – I'm not sure it is – to come up with a framework, we have different conceptualisations of the relationship that could exist between human and artificial intelligence, a bit like the EU AI act that has different ranges of risk (European Parliament,

2023). You could have different types of relationships, and then you could look at each of them and think, 'Well, what are the implications of this for the developer of the AI, and what are the implications of this for the humans and for educators and stakeholders in general?' Then, you could explore each of those. But of course, even if you had that, unless you've educated the population sufficiently, how do they access what that means? Do you see what I mean?

We could do that work, and we could at least come up with something that might be a bit useful. It wouldn't be perfect. That's for sure, but it could be something. But unless people understand enough about AI, they may not be able to make the best use of that. So, it's a real educational challenge. But we have to not go for AI as the autopilot; we have to keep human decision-making in the loop at the appropriate level.

*We have to not go for AI as the autopilot; we have to keep human decision-making in the loop at the appropriate level.*

That means people need to understand what that means, 'Ok, I'm going to let my AI do X, Y and Z'. But what does that really mean? I'm letting it do something, and what am I still making decisions about? It's really challenging. But at least this imagination with a vengeance is opening up the possibility of those conversations in a way that we couldn't have had over a year ago or so.

**ST:** ChatGPT has made particular headlines when it comes to assessment in higher education. It has been said that the essay is dead, and the integrity of open-book online exams has become questionable. We may arrive at a situation where the lecturer creates the assessment using ChatGPT, then the students write the assessment using ChatGPT, and then the lecturer marks the assessment using ChatGPT, and nobody learns anything anymore (Popenici et al., 2023). To what extent does generative AI impact assessment? Is there a legitimate use of generative AI by students and teachers? Does AI call for a reassessment of how we should set assessments? Should we try to make assessments more authentic? Should we try to have more assessments that attempt to be 'AI proof'? Could you please also discuss the potential of AI in revolutionising traditional assessment methods, particularly in providing real-time, personalised feedback to learners?

**RL:** Assessment is a key area of impact. But I think the key question for people when we think about assessment is: How are we going to change our assessment methods so that they actually assess what we need to assess? In the book "Machine learning and human intelligence" (Luckin, 2018), I say that we treasure what we measure, and assessment is our measurement tool. At the moment, in most cases, I don't think we measure the right things, and therefore, we don't treasure the right things.

*The key question is: How are we going to change our assessment methods so that they actually assess what we need to assess? We treasure what we measure, and assessment is our measurement tool.*

*At the moment, in most cases, I don't think we measure the right things, and therefore, we don't treasure the right things.*

If we were measuring more sophisticated thinking skills, then the assessments probably wouldn't be as accessible to something like ChatGPT. But whatever! We don't know how these models are going to develop. So, the key imperative is to think about redesigning assessment and revisit the basics of why we are assessing and what it is we want to assess. We have to move on from assessing memorisation, which is a lot of what we do in the UK. It's perfectly obvious that any quite basic large language model can do pretty well at exams that test memorisation. If they've been trained on the right data, they'll remember it. They might get some things wrong, but they'll still do well.

There's an imperative for us to think carefully about what it is we want to assess and then how we're going to assess it. On some occasions, that assessment may involve the use of AI. Why not? I believe it should be seen like this: If this AI is here to make us more intelligent, then stopping people from using it is a very backward step. So, how do we make the assessments that even when you're using AI, you're still assessing the student's capability, knowledge, understanding, whatever it is you're trying to tap into, to measure accurately?

It's the biggest job. I was recently involved in a roundtable organised by an examination body, trying to explore this very question. What is it we should be assessing? How should we be assessing? How could we use AI in assessment? Of course, that's another element of AI, whether you're using it for automated grading or to provide automatic feedback, perhaps in a formative assessment context, or whether you are using it to analyse and measure student progress as they learn. AI is definitely a powerful tool that can be used as part of the assessment process itself. It's also a tool that students can transparently be encouraged to use as part of the assessment. But before you can get to that stage, you've still got to do the nitty gritty on what it is we're trying to assess and why we want to measure this. Why does this matter?

At the moment, rigour has got in the way of the value of the things that we measure. We think, 'As long as the measurement is rigorous, it's ok'. But actually, now it's not. So, we need to do huge amounts of thinking. That also relates to the question about the relationship we want to have between humans and artificial intelligence. Because if we want to look at the human side of the relationship in a particular way, how do we assess the extent to which we're achieving that successfully?

AI is not going to go away, and it's, without question, radically transforming the workplace. Just look at the legal profession. It's turning cartwheels or somersaults at the moment. Perhaps more accurately, these AI tools can do much of what the legal profession was doing. Of course, not always flawlessly, but I think we owe it to our students to focus on being much clearer about what we want to measure through our assessment systems and why. Then,

we can start thinking about how and what the role of AI might be in that as part of the assessment or as part of the measurement.

Those conversations need to be informed by the kind of relationship we believe society would benefit from between human and artificial intelligence. But the big elephant in the room is who is going to be responsible for that? I remember attending an event in the UK many years ago, which was organised by a body called Becta, which doesn't exist anymore. It was a non-departmental government body that oversaw the way that educational technology was used in the UK and developed framework support, etc. They were organising a whole set of policy workshops. I went to one of these, and there are various policymakers, big tech representatives, educational body representatives and quite a lot of civil servants. At the end of one of these sessions, we were asked the question that was about changing the education system: Who is responsible? Is it the educators and the educational institutions, or is it the policymakers and government, or is it society as a whole? The majority of the room, not myself included, voted for society as a whole because nobody wanted to own it. [All laugh.]

*Who is responsible? Is it the educators and the educational institutions, or is it the policymakers and government, or is it society as a whole? The majority of the room, not myself included, voted for society as a whole because nobody wanted to own it.*

And therein lies a huge challenge. Is it the examination bodies? Is it the government? Is it the schools? It's got to be a cross-sector, multi-stakeholder engagement. But who's going to bring it together? And how do we get it right? A really big challenge!

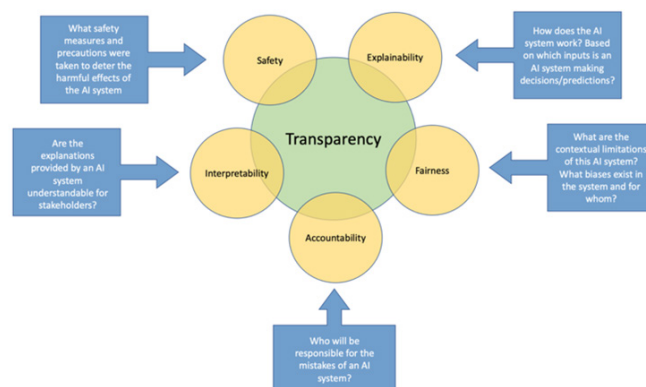


Figure 9: Transparency in relation to other dimensions of ethical AI (Chaudhry et al., 2022).

**JR:** We are now moving on from assessment to teaching and learning, which are, of course, intrinsically intertwined. How can we change our approaches to teaching and learning in light of generative AI? ChatGPT and other chatbots (like Bing Chat, Bard – now called Gemini (and hence no longer insulting to Shakespeare) – and Ernie) are the latest shiny things in the long history of AI in education. It may be too

early to determine ChatGPT's (and GPT-4's) place in that history, but what are your preliminary thoughts? In your view, how should higher education institutions deal with generative AI?

**RL:** The assessment piece is obviously significant, but we've talked about that. 'Strategically' is my answer. I don't think you should separate generative AI from AI in general. Obviously, it has different capabilities, and that needs to be recognised. However, every higher education institution needs a good AI strategy. That should be built on their existing vision, and that institutional vision may need to change. It depends on what it is. You might want to revisit it. But if you have a vision that's about educating a wide demographic of students in a fair and equitable way, there's no reason that would change just because we have AI. The way you do it might change, but the existence of that vision doesn't necessarily change. Then, strategically, how are you going to leverage AI to help you achieve that vision more impactfully, way faster for less money, whatever particular metric you want to use in that context? How's it going to help you achieve that vision? That's what it does come down to.

It's complex, obviously, and I'm going to oversimplify it, but one key element is governance and ethics. You need to put in place at the highest level some governance structure that's specific to AI. Maybe it's a high-level committee. Maybe you even invite external people to sit on it to help you make sure you keep abreast. Maybe some of the tech companies get involved. Maybe there's a sector-wide one. There needs to be some high-level governance and ethics committee. Then, obviously, each institution needs to have different policies that would help the people within that organisation understand how to navigate the boundaries that the governance and ethics committee is setting for that institution.

What's allowable? What are we encouraging? What's not? Why? But then, what are the policies that help people on the ground – students, teachers, department heads, university leaders, whatever – to understand how to put that governance and ethics framework into practice? I would encourage people to think of this as an iterative learning cycle. You need to look across the institution and think about areas where perhaps you're facing particular challenges, and they may not all be in the teaching and learning space. Some of them might be in the back office. You need to look to see whether a particular challenge or activity you need to get done – better, faster or cheaper (whatever it is) – could be enhanced by AI.

This is this idea of purpose-driven AI, and once you've identified some use cases that are driven by purposes, what are the kinds of technologies that you might use for that? We've got a template that we use. What's the purpose of this? What's the AI going to help with? Why might it bring benefit? What are the ethical implications? What are the staff capability implications? What are the technology and data implications? What are the practical implications of actually implementing this? How are you going to collect data and evidence about whether it's actually doing what you want? How are you going to learn as an organisation for that? So, you test a few things out, and if they look like

they've got legs, pilot them, but always look to see how that helps you or is constrained by your existing technology, data, and existing staff capability.

So, obviously, those initial use cases are going to be constrained by your existing technology and human capability infrastructure. But each time it is rolled out, there's learning to be had. So, how are you going to make sure that that learning is passed on to everybody who can learn from that particular use case? How are you going to learn about ways in which your technology and data infrastructure might need to change because of this technology being rolled out, and then gradually, as you iteratively test out these different AI, some of them you will chuck out very quickly? Some of them you'll take to pilot and then chuck out, and some you'll take to pilot and then roll out. But by the time you get to rolling out, you'll have a really good sense of what that AI is meant to achieve, what the implications of trying to implement it are, and what you need to have in place in order for it to work. How, fundamentally, are you going to know whether it's achieving what you wanted to achieve?

So maybe, for example, you might have a use case around automated feedback. Let's be honest. In universities, students often don't get feedback in a timely manner. It's improved, obviously, but often they don't, and AI can actually be quite effective at providing feedback. You may not want it to do all of that. You may want it to be overseen by a human. That's something obviously we need to consider. However, various commercially available applications can help provide feedback for students. Feedback is also a really expensive thing to provide, so there's a real cost implication. So, you might decide that's a use case. Look for your tool. The benefits you're hoping to achieve are faster feedback to students, thus more effective for their learning, reduced workload for lecturers, more information for lecturers (because you're able to analyse all of that automated feedback), and saving costs. There might be five elements of what you're looking for, so then you need to have a framework within which you're going to say, 'Ok, yes or no' if it didn't pass our threshold of satisfaction for that particular area.

So, you have these four main elements: (1) governance and ethics, (2) iteratively rolled-out use cases, (3) the technology and data infrastructure piece and (4) the staff capability piece. They are all highly interconnected, but I think it's always helpful to break things down when you're facing a very complex situation. One thing I learned very early in my AI career studies is that a lot of AI at that time was about understanding a problem. So, this is trying to simplify what's a very complex situation.

So, it's ok to start with these four boxes - the governance and ethics, the use case that is being iteratively rolled out across different areas of the institution, technology and data infrastructure, and staff capability. But that iterative process brings other people in, as per the Golden Triangle. So as I have said, with that governance and ethics piece looking for externals, you can help the university understand more, perhaps, about where the technology is going or perhaps about the law. The Golden Triangle pieces are at the heart of helping educators to be part of that iterative testing and

learning.

## Integrating AI into inclusive education

**MG:** In your view, how can AI be effectively integrated into inclusive education to support learners with diverse needs and abilities? I ask that out of interest in what you mentioned earlier about transparency not necessarily being transparent. Introducing new tools that may not be understood by the end user, or where there is a disparity in the existing skills of class members to adapt to a novel technology, could unintentionally create learning environments that disadvantage certain segments of our student population. What are the key considerations for ensuring AI technologies are accessible and beneficial to all students, regardless of their learning differences?

**RL:** It's a hard question, and there's also an added extra dilemma. There's a minority of students who ethically object to using AI because of its potential damage to the climate. So, there's a real equality issue there. Do you respect their quite heartfelt, passionate views? If you do, then they're not going to have the same access. So, it's definitely a dilemma.

Two of the key features the AI brings to the party are the ability to be adaptive and the ability to behave autonomously, though, as we've discussed, the extent to which you want your AI to be autonomous is something that needs to be thought about carefully. Because an AI works by analysing its environment and then, based on that analysis, the way it's been built and trained, what it's meant to achieve, and what its goals are, it behaves in a particular way. So, if it's a grading algorithm, the environment is the assessment materials it is processing. It's been trained on thousands or millions of previous examples and various other artefacts like rubrics, etc. The goal is to produce some feedback. So, we know that's how it works. Of course, it's adaptive, so the feedback it gives to one student won't be the same as the feedback it potentially gives to another, even if their assignments are identical, although that's an interesting question.

So, in the one sense, it brings that ability to be adaptive and to behave with a certain amount of autonomy that should, in principle, mean that no matter what a student's ability, there should be an AI that can help them and meet them where they are. That's part of the dream. It's this notion of really quite hyper-personalised learning where you have an AI that really gets to know the student. Hopefully, the student gets to know themselves as well, but that's the sort of dream of the AI that enables you to meet any user's needs – because of this adaptivity and autonomy. Of course, the other side of that promise is this dilemma that you really are right to identify: there are issues about accessibility.

We're already seeing it, and it's so true. If you pay for one of these models, they're so much better than the ones that you get free. So, it is an immediate equality issue. The issue that you already highlighted about whether people understand nothing anyway in the first place to be able to access the technologies is a real core accessibility issue. Then you've got this ethical dilemma about whether people passionately feel these technologies shouldn't be used and whether we

need to respect that. On one hand, it brings the potential to really help address challenges of access and equality, and on the other, it brings us a whole new raft of ways in which we can get it wrong.

## AI, teacher education and professional development

**ST:** With the rapid advancement of AI in education, what changes do you foresee in teacher education and professional development programmes? How should teacher training evolve to equip educators with the necessary skills to integrate AI tools effectively in their teaching practices?

**RL:** Again, this is such a core part of the puzzle. They have to change to take it into account. I don't foresee a rapid change, unfortunately. At least it'll be patchy. Some countries will be much further ahead than others, so it's not universally the case. But history teaches us that changes to teacher education and professional development can be very slow. And here we're faced with a technology that's moving very quickly. That is a really big challenge for the education system. So, potentially, the best way of dealing with it is actually through continuous professional development. Of course, you need to look into initial teacher training. But it's not just initial teacher training. It's recognising that this is going to be something that people are going to have to learn about, not just now, but next year, the year after, and the year after that. There is a need to design for continuous professional development to support teachers as the technology changes, but they need to get that base understanding first, and we're not even there yet in most cases.

*History teaches us that changes to teacher education and professional development can be very slow.*

So, the pressing need is to help educators get a basic understanding of AI so that they're what I would call 'AI-ready'. Then, you can start looking at how you develop their AI readiness into different skills, abilities, and capabilities. Maybe some people in an organisation specialise in one type of AI or one where AI is used in a particular channel, and that will be up to organisations to decide. But it also needs to be part of that overarching strategy that an institution needs to have. There are these constant cycles of learning.

Two further very practical things. It's essential that senior management are very vocal about their support for training because they're always making a difference in how well it works. The teachers are given time and recognised for doing the training so that it's something that they feel good about, that they feel they want to do. We really need them to do it because they're being faced with students who are using these technologies. Anecdotally, I was talking to an organisation recently about some research they've been doing on young people where some school-age learners are. In fact, the majority see AI as the solution to all their problems. That's a real worry.

*The majority see AI as the solution to all their problems. That's a real worry.*

If we don't have educators who have a basic understanding to help them navigate that path better than we've done with social media, then we're in a difficult position. After a basic understanding, continuous professional development needs to be built into each institutional strategy. As you iteratively roll out the AI, everybody's learning, and people are open and honest and feel ok about being honest. It's a bit like dealing with phishing emails. You need people to tell you when they've inadvertently opened that attachment. You don't want them to feel scared because you need to know they've opened that attachment. You need people who feel that if they get something a bit wrong with AI, they're actually going to tell you about it because you need to know! [All laugh.]

**JR:** I've clicked on those phishing emails before.

**RL:** Yeah, everybody has.

**ST:** Considering the importance of lifelong learning in today's dynamic world, how do you see AI contributing to adult education and continuous professional development? What are the opportunities and challenges in this area?

**RL:** This goes beyond the education profession. World Economic Forum (2023) data show the expectation that the majority of the workforce isn't very well prepared for the future. There's an imperative there for companies to prioritise ongoing training, which they often don't do. But they're going to be impacted by AI whether they like it or not because their customers are going to use it even if their staff don't. So, it's a global challenge. Few businesses are completely unlikely to be impacted at all by AI. Even if you're running a corner shop, for example, and you feel that it's all about your personal relationship with your customers who come in face to face, and you meet them, and you sell them, and you build relationships, some of them will be using AI. In general, there is an ongoing need for lifelong learning and professional development and for that to be prioritised.

## Evaluating the effectiveness of AIEd

**MG:** You mentioned earlier that, under the Golden Triangle, you should be sure that you are able to evaluate the effectiveness of the things that you implement. So the question is, what methodologies or frameworks do you recommend for evaluating the effectiveness of AI-based educational tools and interventions? How can educators and institutions assess the impact of these technologies on learning outcomes to ensure that they are pushing student learning in a positive direction?

**RL:** We've got to be open and innovative in the way that we approach the evaluation. This isn't a space for randomised control trials, for example, not for a long time anyway, and for many reasons. We need to think about methodologies that enable us to collect quantitative and qualitative data to understand the experience because that is supremely important. In this instance, we can also use AI to help us with that data analysis. So, we can do intelligent analytics on the data that we collect – not just the quantitative data, but also the qualitative data. So, we have the potential to be

quite revolutionary in the way that we evaluate the extent to which a particular AI is or is not supporting the student, the teacher, the parent, or whoever is focusing on it in the way that we want.

But again, that's a question of access because most people don't understand enough about the potential of intelligent analytics and wouldn't know how to necessarily do it. So, we have to start simple but look at ways in which you can mix the quantitative and qualitative data that you collect in a very formative way. When you're doing that initial test, which you could think about as an MVP [minimally viable process], is the technology even worth considering in any detail? What's the MVP that we want to test here? 'Ok, it's passed that threshold'. So, there are different ways of evaluating for different purposes.

'Ok, we're going to pilot it.' What do we want the pilot to achieve? Then, I think using the logic model approach is a great framework for doing this. But to be thinking all the time, how much more sophisticated are the data that's collected through that logic model framework? It would be the subject of some intelligent analytics in the future so that when you get to the stage of actually rolling out an AI across a part of the institution, or even the whole institution, you've got the right data being collected. You've got the analytics in place to make sure that you can learn the things you really want to learn about. So, it's in stages, but the logic model is a great framework to use for that.

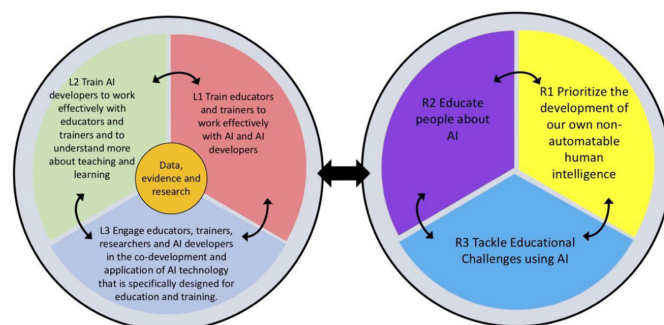


Figure 10: A co-design framework for AI to be used in education and training (Luckin & Cukurova, 2019).

## AI and the role of the teacher

**MG:** Oftentimes, for us working in the private sector in particular, there is a need to justify financially not just the implementation of the technology but also the degree to which these initiatives impact student outcomes, progression, and the student experience. If we look at it from the perspective of the technology, we often see it changing in advance of important considerations regarding its implementation. The question is, how do we keep an eye on balancing the use of these AI and digital tools while still retaining the human-centricity of the learning experience? How do we support students with an AI-curated learning environment without robbing them of the auxiliary benefits of having human educators?



**RL:** Several years ago, I was asked what my dystopian and utopian views of the future of AI in education were, and at that time, I thought what I was saying about the dystopian view was just a nightmare that wouldn't happen. The dystopian view would be that you would have students, generally from poorer backgrounds, being educated mainly by AI systems that were hyper-personalised and adapted to their needs. So, lots of technology interacting with the students, probably some bouncers to ensure that students behaved, but very little human interaction because, after all, that's really expensive, isn't it? Once you've invested in the AI systems, they can work 24/7. They don't need holidays. They don't go on strike. They rarely break, and they get better.

*The dystopian view would be that you would have students, generally from poorer backgrounds, being educated mainly by AI systems that were hyper-personalised and adapted to their needs. So, lots of technology interacting with the students, probably some bouncers to ensure that students behaved, but very little human interaction because, after all, that's really expensive.*

The utopian view, which, obviously, would have been for the more privileged students, was one where the human was much more in evidence. The learning is very human-driven, with the technology very much in the background, orchestrated by the student and the teacher working together. Yes, it is hyper-personalised but very much driven by the humans in the room, with the students gaining a really deep understanding of themselves as learners, of the data that could help them understand themselves, of how to control them, of the ethical implications, the way to benefit from it – but that being in the background.

*The utopian view, which would have been for the more privileged students, is hyper-personalised but very much driven by the humans in the room, with the students gaining a really deep understanding of themselves as learners.*

To be honest with you, I am actually really worried that a dystopian perspective is more likely now than it was a few years ago, and I'm worried about it because I see the dilemma. If I am in a country where there are millions of children who get no education and a technology company comes along to me and says, 'Don't worry, we can give you a hyper-personalised adaptive assistant for each child'. You're not going to say 'no', are you? Because at the moment, I've got nothing, and this would give them something, but to what extent does that then become the thin end of the wedge?

You would be comparing it to nothing that the students would otherwise be learning. If you're able to demonstrate, as I'm sure you would, to what extent do other people think, 'Oh, we could save some money here!?' 'We're financially

challenged.' This comes back to that core question that we've talked about several times, which is, 'What do we want the relationship between the human and the artificial intelligence to be?' Where is the sweet spot where both benefit? We want the AI to get better at giving us what we need. Maybe it's not what we want, and we need humans to get better at making sure the AI gives them what they need. Not necessarily what they want, so that relationship is still core.

## The future of work

**JR:** What will generative AI do to graduate and academic employment? Daniel Susskind (2020) recently wrote a book titled "A world without work". You appear to adopt a more optimistic view (for instance, in Luckin et al., 2016). I believe you are saying that AIED will help people continue to be employable. What are important skills and competencies for graduates to become and remain employable? How do you see the future of academic work in light of generative AI?

**RL:** Generative AI has the potential to automate some routine cognitive tasks currently performed by humans, but I believe there will still be a strong need for our unique human skills and competencies among graduates and academics. Key skills for ongoing employability include:

- Creativity and innovation: Coming up with novel ideas, making connections between disciplines, and designing innovative solutions. AI can enhance but not fully replace human creativity.
- Social and emotional intelligence: Understanding emotional cues, building relationships, empathy, communication, and collaboration. These distinctly human abilities will become even more valuable.
- Complex problem solving and critical thinking: Formulating, analysing, synthesising, and solving new complex problems that require reasoning, judgment and strategy development. AI can augment but not replicate advanced human cognition.
- Adaptability and self-management: The ability to continuously upskill, learn new things quickly, and manage one's own learning and career. This lifelong learning mindset enables graduates to evolve alongside technological disruptions.

Within academia, generative AI will likely enhance academic productivity through automated basic research and writing assistance. But uniquely human skills like conceptualisation, creativity, complex critical analysis, judgment, social perceptiveness, and wisdom will become even more valuable among academics. AI will augment technical aspects, allowing more focus on the interpretative and social dimensions of scholarship. Lifelong learning for faculty to continually advance their own expertise and teaching ability will also grow in importance.

## Artificial General Intelligence and ethics

**JR:** We are now looking a bit more into the future. Nick Bostrom, a philosopher at Oxford University, has written a book about superintelligence (2014). He is cautioning that after computers have achieved Artificial General Intelligence (AGI – which essentially means that they can think and act like humans), superintelligence may be close. This would mean that machines would be potentially and exponentially more intelligent than us humans. One possible outcome would be an extinction event for humanity. A more benevolent one, perhaps, would be that computers keep us as pets. Movies like *The Terminator* and *The Matrix* immediately come to mind. What are your thoughts on this? How do you view the ethical responsibilities of AI researchers?

**RL:** The prospect of super-intelligent AI that surpasses human capabilities gives me great pause for thought and some significant worries. I don't believe there will be an AI that can do everything a human can do, and all of it will be better than a human. But I do believe that with the right emphasis on ethical foresight and safeguards, we can develop AI that enhances rather than endangers humanity.

As an AI researcher focused on education, I feel a profound responsibility to ensure the work I'm involved with broadly benefits society and empowers humans rather than replaces them. Key principles we must embed into AI systems are transparency, accountability, respect for human dignity and agency, and optimising for the social good rather than solely efficiency or profit motives. Robust policy frameworks on areas like data rights, preventing bias, and adherence to human rights also need to be implemented alongside the technology as it advances. Multi-stakeholder consultation is key.

*As an AI researcher focused on education, I feel a profound responsibility to ensure the work I'm involved with broadly benefits society and empowers humans rather than replaces them.*

Rather than AI systems that control critical decisions autonomously, I believe human-AI collaboration is crucial – designing intelligent assistants that enhance human judgment, creativity and well-being. Systems focused on personalised learning over standardised outcomes respect learner agency, too. The goal must be expanding human potential – economically, creatively, socially and culturally – not limiting it. With ethical AI guardrails and human-centric design principles in place from the outset, I am hopeful we can positively shape this technological frontier for the betterment of humanity as a whole.

*The goal must be expanding human potential – economically, creatively, socially and culturally – not limiting it.*

**MG:** In "Intelligence Unleashed: An argument for AI in education" (Luckin et al., 2016), you present a vision for AI in education. Looking forward, what emerging trends or technologies in AI do you believe will be most transformative

for education in the next decade? Finally, could you share with us your future plans and any additional topics or advice you'd like to offer to upcoming researchers in the field of AI and education? And is there anything else we did not cover that you would still like to discuss?

## The future of AIEd

**RL:** The continuing rapid pace of progress in AI currently is exciting. Areas like natural language processing (NLP), personalisation via machine learning (ML), multi-modal interaction, and human-AI collaboration hold particular promise over the next decade. For example, NLP could enable richer dialogue and feedback between the learner and AI tutor. Personalised machine learning could allow more targeted content, guidance and support customised to each student's strengths and weaknesses. Multi-modal AI that incorporates speech, vision, and haptics, along with language understanding, could make interactions more intuitive and assistive for a wider range of learning needs. And improved techniques for human-AI complementarity will be able to amplify learner potential. As long as we get the ethics, safeguarding, and guidance correct, these are just some of the significant benefits we can expect.

I envision a future with AI-empowered personalised lifelong learning companions that can adapt to each individual, motivate and guide self-directed growth – companions that are firmly in the control of the user and there to support them to become ever more intelligent and ever more knowledgeable about themselves as a learner. Learners should be creators and partners alongside their AI tutors, definitely not just consumers of content. Educational equity can expand tremendously if we get the guidance and regulation of AI progress right.

As someone pursuing research in this exciting field, I will focus first on the key human challenges to address rather than leading with technological capabilities. I will always explore the ethics, the potential for inclusion and the human development needs from the outset. As always, I will continue to collaborate actively with educators, learners, and policymakers to ensure real-world relevance and responsible progress as we try to ensure that we humans become ever more intelligent – especially in uniquely human ways. There are still so many open and exciting questions that merit creative exploration around peer learning, meta-cognition skill building, transferable competencies, and the interplay between AI and quality human teaching, for example. I think I am likely to continue to be very busy!

**JR:** We are very grateful to you for this outstanding interview.

## References

- Bostrom, N. (2014). *Superintelligence. Paths, dangers, strategies*. Oxford University Press.
- Brookfield, S. D., Rudolph, J., & Tan, S. (2023). *Teaching well. Understanding key dynamics of learning-centered classrooms*.

Routledge. <https://doi.org/10.4324/9781003447467>

Brooks, T., Peebles, B., Holmes, C., DePue, W., Guo, Y., Jing, L., Schnurr, D., Taylor, J., Luhman, T., Luhman, E., Ng, C., Wang, R., & Ramesh, A. (2024, February 15). *Video generation models as world simulators*. OpenAI. <https://openai.com/research/video-generation-models-as-world-simulators>

Chaudhry, M. A., Cukurova, M., & Luckin, R. (2022, July). A transparency index framework for AI in education. In *International conference on artificial intelligence in education* (pp. 195-198). Springer International Publishing. [https://doi.org/10.1007/978-3-031-11647-6\\_33](https://doi.org/10.1007/978-3-031-11647-6_33)

Cukurova, M., Luckin, R., & Clark-Wilson, A. (2019). Creating the golden triangle of evidence-informed education technology with EDUCATE. *British Journal of Educational Technology*, 50(2), 490-504. <https://doi.org/10.1111/bjet.12727>

Erricker, D. K. (2023). *Lesson 21: Technology, innovation and system change* [Presentation]. <https://www.cite.hku.hk/wp-content/uploads/2023/10/20231030keynote2.pdf>

European Parliament. (2023, December 19). *EU AI Act: First regulation on artificial intelligence*. <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>

Hofstadter, D. R. (1999). *Gödel, Escher, Bach: An eternal golden braid* (20th anniversary ed.). Basic Books.

Luckin, R. (2010). *Re-designing learning contexts: Technology-rich, learner-centred ecologies*. Routledge. <http://dx.doi.org/10.4324/9780203854754>

Luckin, R. (2018). *Machine learning and human intelligence: The future of education for the 21st century*. UCL Institute of Education Press.

Luckin, R. (2020, August 4). *Is education AI-ready? Turing lecture* [PowerPoint presentation.] [https://www.turing.ac.uk/sites/default/files/2020-08/turing\\_lecture\\_-\\_is\\_education\\_ai\\_ready\\_pdf.pdf](https://www.turing.ac.uk/sites/default/files/2020-08/turing_lecture_-_is_education_ai_ready_pdf.pdf)

Luckin, R., Bligh, B., Manches, A., Ainsworth, S., Crook, C., & Noss, R. (2012). *Decoding learning: The proof, promise and potential of digital education*. Nesta. [https://www.researchgate.net/publication/269111789\\_Decoding\\_Learning\\_The\\_Proof\\_Promise\\_and\\_Potential\\_of\\_Digital\\_Education](https://www.researchgate.net/publication/269111789_Decoding_Learning_The_Proof_Promise_and_Potential_of_Digital_Education)

Luckin, R., & Cukurova, M. (2019). Designing educational technologies in the age of AI: A learning sciences-driven approach. *British Journal of Educational Technology*, 50(6), 2824-2838. <https://doi.org/10.1111/bjet.12861>

Luckin, R., & Cukurova, M. (2023). Small-scale commercialisation: The golden triangle of AI EdTech. In B. du Boulay, A. Mitrovic, & K. Yacef (Eds.), *Handbook of artificial intelligence in education* (pp.537-552). Edward Elgar Publishing. <http://dx.doi.org/10.4337/9781800375413.00036>

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, 3, 100076. <https://doi.org/10.1016/j.caeai.2022.100076>

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. <https://discovery.ucl.ac.uk/id/eprint/1475756>

Popenici, S., Rudolph, J., Tan, S., & Tan, S. (2023). A critical perspective on generative AI and learning futures. An interview with Stefan Popenici. *Journal of Applied Learning and Teaching*, 6(2), 311-331. <http://dx.doi.org/10.37074/jalt.2023.6.2.5>

Shah, S. (2024, February 14). *ChatGPT 5 release date: What we know about OpenAI's next chatbot*. The Standard. <https://www.standard.co.uk/news/tech/chatgpt-5-release-date-details-openai-chatbot-b1130369.html>

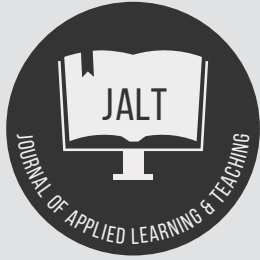
Susskind, D. (2020). *A world without work: Technology, automation and how we should respond*. Penguin.

Team EdTechX. (2019, November 14). *UCL EDUCATE demo day 2019*. Medium. [https://medium.com/@team\\_edtechx/ucl-educate-demo-day-2019-7856ff1fd2ea](https://medium.com/@team_edtechx/ucl-educate-demo-day-2019-7856ff1fd2ea)

Weatherby, K., Clark-Wilson, A., Cukurova, M., & Luckin, R. (2022). The importance of boundary objects in industry-academia collaborations to support evidencing the efficacy of educational technology. *TechTrends*, 66(5), 784-797. <https://doi.org/10.1007/s11528-022-00705-0>

World Economic Forum. (2023). *The future of jobs report*. <https://www.weforum.org/publications/the-future-of-jobs-report-2023/>

Copyright: © 2024. Rose Luckin, Jürgen Rudolph, Martin Grünert and Shannon Tan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



## 'Failing well' in teaching about race, racism and white supremacy. An interview with Stephen Brookfield

Stephen D. Brookfield<sup>A</sup>

A

*Adjunct Professor at Teachers College, Columbia University (New York) and Emeritus Professor at the University of St. Thomas (Minneapolis-St. Paul)*

Jürgen Rudolph<sup>B</sup>

B

*Director of Research & Learning Innovation, Kaplan Higher Education Academy*

Shannon Tan<sup>C</sup>

C

*Research Executive, Kaplan Higher Education Academy*

### Keywords

Black Lives Matter;  
Critical Race Theory (CRT);  
critical theory;  
higher education;  
intersectionality;  
race;  
racism;  
teacher training;  
white privilege;  
white supremacy.

### Abstract

Since embarking on his educational journey in 1970, Professor Stephen Brookfield has worked across various international settings, including England, Canada, Australia, and the United States. His experience spans a diverse range of environments, from adult and community education to prestigious higher education institutions like Harvard University and Columbia University. Central to his mission is aiding adults in critically examining prevailing ideologies they have absorbed. To advance this goal, Professor Brookfield has authored, co-authored, or edited 21 books encompassing topics such as adult learning, teaching methodologies, critical thinking, discussion techniques, critical theory, and anti-racist teaching.

### Correspondence

Jurgen.Rudolph@kaplan.com <sup>B</sup>

Expanding upon our previous dialogues with Stephen Brookfield in the *Journal of Applied Learning & Teaching* (Brookfield et al., 2019, 2022) and complementing the reviews of his recent publications (Rudolph, 2019, 2020, 2022; Waring, 2024), this interview delves deeper into the themes explored in our recent book on *Teaching well* (Brookfield et al., 2024). This extensive conversation significantly elaborates on Chapter 9 of the book (Brookfield et al., 2024) and investigates the intricate, emotionally charged, and political project of teaching about race.

### Article Info

Received 6 June 2023  
Received in revised form 27 February 2024  
Accepted 29 February 2024  
Available online 11 March 2024

DOI: <https://doi.org/10.37074/jalt.2024.7.1.23>

In this expansive discussion, we explore Stephen Brookfield's personal evolution from harbouring racist beliefs in his youth to embracing and contributing to Critical Race Theory (CRT), a journey marked by a decade of introspection and scholarly exploration, culminating in several key publications (Sheared et al., 2010; Brookfield & Associates, 2018; Brookfield & Hess, 2021). The conversation illuminates fundamental concepts such as race, racism, and white supremacy, recontextualising racism as a systemic issue rather than an individual failing. Racism is depersonalised and an endemic system of exclusion. We discuss it in the context of an intersectional analysis that acknowledges the interconnectedness of various forms of oppression, including classism, sexism, and ableism. A significant focus is placed on racism within the higher education sector. Brookfield shares insights from his extensive experience in conducting antiracist workshops for students, faculty, and organisations. He challenges the notion of the 'good white people' and advocates for a continuous, imperfect journey towards antiracism, where 'failing well' can be regarded as a good outcome.

## Racism and white supremacy during Stephen's childhood and youth

**Jürgen Rudolph (JR):** Teaching race is one of the most complex pedagogic projects, and doing it well requires us to shift our notions of what counts as success quite substantially. Teaching race is an emotional and political project. You wrote: "Because I'm used to seeing myself as un-raced, it took me a long time to realise the truth of an African American co-teacher's comment: 'to students of color everything is seen in racial terms'" (Brookfield, 2017, p. 139). White people appear to have a particular problem with racism, as whiteness "is still widely taken for granted and thus remains invisible" to them (Cunningham, 2010, p. xxvi). Being color-blind is, however, an illusion. In your article "Teaching our own racism" (Brookfield, 2014, p. 90), you discussed your beliefs and attitudes growing up as a white person: "Attitudes and beliefs I picked up in my childhood, adolescence, and young adulthood included Blacks were alternatively lazy, happy, or violent; Pakistanis and Indians were sexually irresponsible having large families; and Gypsies were thieves out to mark your house as an easy target". These kinds of prejudices and stereotypes all sound terribly familiar to me from my own upbringing.

When you were 17, you encountered what is called in critical race theory, a counter-story. You wrote:

I was being beaten up by a gang of white youths (they were 'rockers', I was a 'mod') in an English town one Friday night. A black American serviceman from a nearby USAF [United States Air Force] base crossed the street and broke up the fight telling us 'everybody's got to be cool now'. In my memory, I was on the verge of falling to the floor as the GI [U.S. soldier] intervened to save me from potentially severe injury (Brookfield, 2017, pp. 214-215).

Did the incident lead to any changes in your own beliefs when it comes to race? Could you elaborate on how you experienced and learned racism and white supremacy during your childhood and youth?

**Stephen Brookfield (SB):** That event is easily recalled. It stayed with me for the rest of my life as a clear interruption of this dominant narrative that I learned growing up regarding the different stereotypes attached to different racial identities. Of course, under white supremacy, Whites are cast as the non-violent users of reason and logic and people of color. Specifically – with the anti-blackness part of white supremacy – Blacks are often cast as inherently volatile, unpredictable and with a propensity for violence. So, that whole equation was shattered by that particular event because if he hadn't intervened, I think I would have been hurt much more than I was, which was basically just bruises and cuts that needed to be dressed. There weren't any internal injuries or anything like that that I suffered. That has stayed with me.

It was one of my earliest encounters with race because, at the time, I lived in an English village close to a market town. The village was almost completely white, and the town was overwhelmingly that way. It was a very dramatic illustration

to me that the ideas that I was learning around race were inaccurate. In terms of how I learned these instincts and impulses, behaviors, and actions, I learned it by a process of unconscious internalization – I didn't realize I was learning these things, which is the nature of white supremacy. I never saw anyone in authority who wasn't white. Those in authority were also overwhelmingly male. And the images that I had of people who were to be admired were, through media, in politics and even in sport – because I followed soccer, in particular – were overwhelmingly white, though there were alternatives to that: the Brazilian soccer team was often held up as the pinnacle of the beautiful game. But mostly, my daily life was just soaked in this whiteness.

*Mostly, my daily life was just soaked in this whiteness.*

So, that's how I grew up. I assumed that leaders, in particular, were white, and those that we looked up to were white. The people who exercised power and authority in schools, religion and definitely in politics, and those who were known as successful in business were all white. So, I think that was something that soaked into me at a very unconscious level. I really wasn't aware. A fish can't see the water that they swim in. So, you're not aware of your breathing, the thing you do all the time to keep you alive. It's just the same as white supremacy, and of course, it was supported – I'm sure it was the same in Germany – by media images.



Figure 1. BBC's Black & White Minstrel Show. Not dated (the show ran from 1958 to 1978). Source: The Black and White Minstrel Show (n.d.).

When I was growing up, there was something called the Black and White Minstrel Show on the BBC (see Figure 1). Every Sunday night, we would watch these white male and female singers in blackface, caricaturing the old minstrel shows (see Figure 2). It was a straight reproduction of them.

I remember having Robertson's jam – a famous jam, where there was what they call a Golliwog on it, which was a little blackface boy (see Figures 3 & 4).



Figure 2. The Strobridge Litho Co. originally published this reproduction of a 1900 William H. West minstrel show poster. It shows the transformation from a person of European descent to a caricature of a dark-skinned person of African descent. This image is available from the United States Library of Congress's Prints and Photographs division under the digital ID var.1831.



Figure 3. Robertson's jam's Golliwog (Petcher, 2012).

These images were everywhere. There were also a lot of stereotypical racial jokes that my friends and I would tell without even thinking about it. In friendship networks, I was captain of the soccer team in my school. Whenever we played a match, there were lots of jokes that were told around race. Even in my own family, there were references to different races, not all of them negative. Still, where South Asians from the Indian subcontinent were involved, they



Figure 4. Florence Kate Upton's Golliwog in formal minstrel attire in *The adventures of two Dutch dolls and a 'Golliwogg'* (2023), first published in 1895. Public domain.

were pretty racist.

I had no interaction with anyone other than that GI. Then, when I went to college at 18, that was the first time I really had any sustained conversations with anyone who wasn't white. It's interesting to me that when I was in college, which was from 1967 to 1970, those dim and distant days in the last century [all laugh], I did all the right racial things. At the time, the South Africans still had a strong apartheid regime, and the South African rugby team was touring England. We would go to demonstrate against that at the rugby grounds. I remember being charged by police on horses. I had a black roommate and Pakistani friends. On one level, I had, I guess, antiracist credentials, but as always, for me as a white person, I could really choose when it came to race. I could say, 'Okay, today I need to think about race and talk to my friends about it,' and then on other days, I didn't need to, whereas they had to deal with it every day, of course. I was at a kind of cognitive understanding to some degree from 17 to 18 years old. But I don't think it had any emotional or visceral resonance until much later.

### Confronting historical legacies through anti-racist advocacy

JR: My next question follows up on this. You were born in the UK, a country with a history of colonialism (and its inherent racism and white supremacy that you just illustrated so well). I was born in Germany, a country that will forever be associated with the Nazis, the Holocaust, and other horrible atrocities based on racist beliefs during the Third Reich.

When it comes to the U.S. (where you have spent most of your life), racial injustice began with the original sin of slavery. Even after its hard-fought abolition, it has endured due to white supremacist beliefs and racial discrimination. We believe you became increasingly aware of your own 'race-blindness' in the early 1990s and started to study the

topic extensively during a self-imposed silence on race (in terms of publishing) that lasted a decade. Since 2003, you finally began to write about race, especially in the context of higher education. You wrote in your most recent book:

We were inspired by the Black Lives Matter movement, outraged by the growth of anti-Blackness in the United States, and staggered by the way it became legal to tear immigrant families apart at the U.S. border and imprison children like animals in cages. Each week brought further instances of the slaughter of people of color and the demonization of anyone not of white European descent (Brookfield & Hess, 2021, p. 13).

What motivated you to write, in addition to numerous articles, two books on teaching race and becoming a white antiracist (Brookfield & Associates, 2018; Brookfield & Hess, 2021) and coedit *The handbook of race and adult education* (Sheared et al., 2010)? Was it a sense of a never-ending grave injustice and unfairness?

**SB:** This is a good question to think about. I must have said this somewhere, or you intuited it that I didn't write about this for a while. Like a lot of Whites in the U.S., I felt like it wasn't my place to be writing about race because I really hadn't any sustained experience of being on the receiving end of racism. But as a critical theorist, you're always interested in how permanent inequality becomes seen as normal, natural and just, accepted as the way the world is organized by some natural, universal law. Of course, race was a part of that and I knew all that. So, I have always had an interest, since I was aware of that theory as a young guy, in how things are set up in a society to make these massive discrepancies of power and access seem like common sense.

*As a critical theorist, you're always interested in how permanent inequality becomes seen as normal, natural and just, accepted as the way the world is organized by some natural, universal law. Of course, race was a part of that.*

But what really focused me on race was in the '90s, as you say. I had an experience of working for about ten years with two African-American women who were colleagues of mine in a program I set up – a doctoral program in adult education that I had helped create in Chicago at a university called National Louis University. One-third to a half of the students were African American or people of color, which was unusual for me as I hadn't had that strong a representation in classes I'd taught before. I was in a teaching team of three. My other two colleagues were African-American women: one of whom was Elizabeth Peterson, a critical race theorist, and one of whom was an Afrocentric theorist, and her name was Scipio Colin III. In fact, Elizabeth and Scipio are both co-editors of the *Handbook on race and adult education*, which I was one of the co-editors of (Sheared et al., 2010).

That experience was crucial for me. We were teaching about adult learning and adult education, but a lot of it brought up questions of racial identity. Through Scipio, I was looking

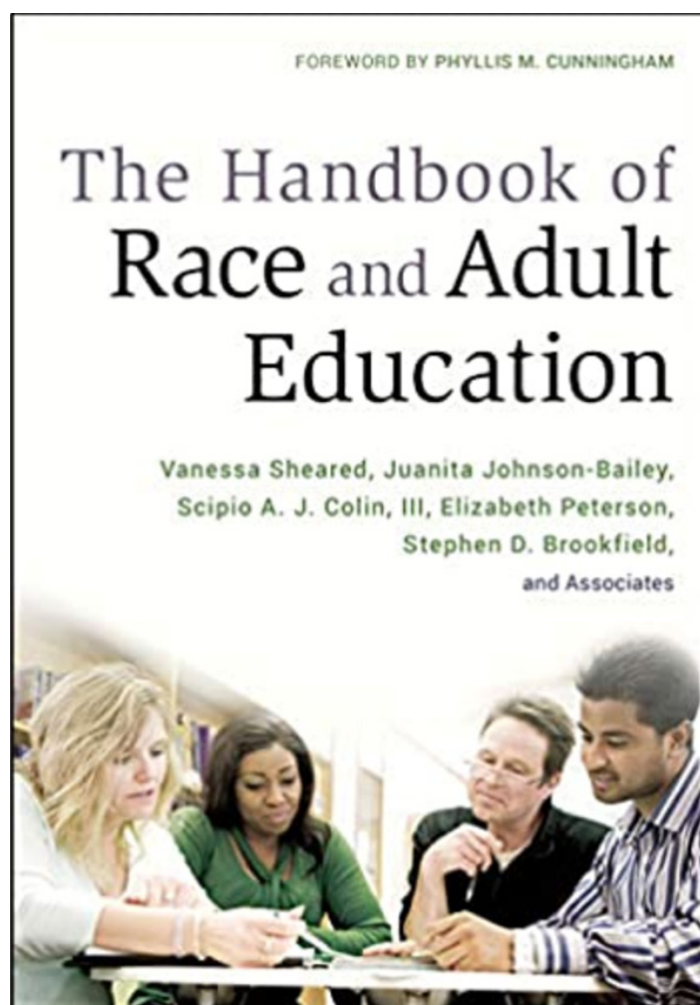


Figure 5: Book cover of Sheared et al.'s *The handbook of race and adult education* (2010).

at this Afrocentric model, which was an alternative to a Eurocentric model, and thinking: how would an Afrocentric model play itself out in terms of learning and in terms of adult educational practice? Elizabeth really introduced me to critical race theory, the works of Derrick Bell (1995) and so on. So, that was a big learning experience for me and just working for a decade with a lot of students of color and those two colleagues of color – plus others, but those two in particular – it was an education in the experience of what it's like to be surrounded by white supremacy: even in Chicago, which is known as a very multiracial city: how they experienced racism all the time, both the students and my colleagues.

I also remember being called out in the early '90s by a colleague of mine, Elizabeth Kasl. She was a colleague of mine first at Teachers College, and then we kept in touch. She moved out to the California Institute of Integral Studies in San Francisco. I remember vividly having coffee one day and her asking me, 'Where is race in your books?' And I said, 'Well, I don't really need to deal with race; I'm talking about learning'. She challenged me on that, and that was a big conversation for me as well.

Also, I had this developing awareness that to be on the right side of history, you need to be trying to work in an anti-racist way. But also, this understanding of white supremacy,

while it was advantaging me as a white person, was also damaging me as a white person. What I mean by that is if you believe at some superficial level in white supremacy, if you buy that myth as it were, you have to live in this state of suspended belief. Because everywhere you see examples that counter white supremacist assertions about white people being calm and reasonable and using logic, and therefore being entrusted with making decisions because they stay objective.

That clearly is so obviously not the case. In some ways, I often think that the four years when Trump was in power certainly deepened white supremacy, but also made it clear to everyone how crazy that idea is, as you looked at what that administration did and how they acted. So, you can't really live believing in white supremacy without doing this cognitive bifurcation and disregarding all this empirical evidence about the insanity of decisions from white people and the craziness of their actions in order to keep white supremacy intact in your head. I don't think many people are able to do that.

On the one hand, you need to be able to escape this schizophrenic state that you're living in. On the other hand, if you really, truly believe that everybody of color is potentially violent and unpredictable and way too emotional and things can erupt at any moment with them, then, as you live in a multiracial world as a white person, you're living in a state of constant fear, uncertainty, and mistrust of people of color around you. Although the world of higher education is overwhelmingly white, it is not completely white, and community colleges, which I think are at the frontlines of higher education in the United States, are much more racially representative of what the United States looks like. I've done much work with community colleges. So, it's unhealthy to live with this constant fear of the other.

When I moved to New York in 1982 to take up a position at Teachers College [Columbia University], we lived right on the edge of Harlem. To go through the day fearful every time I saw a face that wasn't white would have been just horrendous for me. A lot of what I say to white people that I'm working with is – I appeal to a sense of self-interest saying – 'Look, racism has all these social and moral sins attached to it. We need to be fighting it for the sake of having just common humanity, but we also need to be fighting it because it benefits us to fight it as Whites'. I've found that argument – where you use your own self-interest – tends to reach people sometimes in ways that the transformative and moral arguments about the need to combat this inhumane system tend not to.

So, that's a long, rambling answer. But that's some of the motivations behind what got me into this. I decided I couldn't write anything on this until I've got a decade of really thinking about it and reading about it. So, I think my first piece on race came out in 2003, maybe; that's when I started publishing about it.

## Challenging racial narratives: Dissecting racism and white supremacy

**Shannon Tan (ST):** Thank you for your fabulous answer. It was not rambling at all. By the way, it was not intuited that you had this ten-year pause; you wrote that somewhere. With discussions on race being often emotionally charged, some of the key terms (especially racism and white supremacy) are bound to be understood differently and contested. For instance, there is the prominent fairy tale that "deep racism doesn't exist anymore, that any Black person who works hard enough can become economically self-sufficient, that women have gained equality with men to the extent that White males are now the victim of minorities and domineering women, and that those who are poor and unemployed are in that state by choice" (Brookfield, 2005, p. 331). How can we define some of the key terms, such as racism and white supremacy?

In *Becoming a white antiracist*, you write that race "is not real" in the sense that "race as a biologically determined category is a complete illusion", but "racism is very real" (Brookfield & Hess, 2021, p. 30). In Sheared et al. (2010, p. 2), you define racism as follows: "Racism exists when one racial group has power and authority over another racial group because of beliefs about race". In the same book chapter, you say that racism is "the ugly operationalization of the ideology of White supremacy" (Sheared et al., 2010, p. 15). In *Teaching race* (Brookfield, 2018, p. 2), you define racism as a "system of beliefs and practices that are embedded in the institutions we move through as individuals and routinized in the conventions of everyday lives". In the same book, Pamela Barnett (2018, p. 123) posits that racism is a system of beliefs and practices -

in which public policies, institutional practices, cultural representations, and other norms work in various, often reinforcing ways to perpetuate racial group inequity. It identifies dimensions of our history and culture that have allowed privileges associated with 'whiteness' and disadvantages associated with 'color' to endure and adapt over time. Structural racism is not something that a few people or institutions choose to practice. Instead, it has been a feature of the social, economic and political systems in which we all exist.

Having observed that racism is structural, you advocate a systemic understanding of it:

If racism is seen as an act of individual choice or individual sin, then acknowledging one's racism becomes mixed up with viewing oneself as an evil purveyor of hatred and bigotry. But if students become used to seeing racism as a systemic phenomenon, an ideology that is embedded and routinized in practices, habits, and structures that we are exposed to from an early age, then it become[s] obvious that for Whites not to have learned racism is impossible. So, constantly clarifying the systemic nature of racism is an important teaching act (Colin et al., 2010, p. 365).

Racism must be one of the most sensitive topics ever. To cite a quote from *Becoming a white antiracist*:



[B]eing called a racist is considered a *very bad thing*... 'the worst thing to happen to anybody anywhere' [Oluo, 2018, p. 213]... For those who think of themselves as good, color-blind whites, it's the ultimate insult because it's usually applied to people who do overt violence against people of color, commit blatantly destructive acts, and use hate speech" (Brookfield & Hess, 2021, p. 89; emphasis in original).

The way you use the term 'white supremacy' does not refer so much to obvious examples such as the Ku Klux Klan (KKK), Aryan Nations and other extreme white nationalist terrorist groups, but rather "the idea that whites, because of their superior intellect and reasoning power, should be in control of decision-making for society as a whole" (Brookfield, 2018, p. 4). You perceive white supremacy as the all-pervasive "philosophical foundation of racism" (Brookfield, 2018, p. 4). This being a very contentious term, you also use alternate terms such as 'white advantage', 'white superiority', 'white privilege', 'white normativity', or 'white racial frame' (Brookfield & Hess, 2021, p. 32).

Can racism be purely understood in terms of skin colour? Within orientalism, the so-called 'Orientals' (whose skin colour could be as fair as that of whites) were pejoratively characterised as "backward, degenerate, uncivilized, and retarded" in order to be subjected and colonised by white supremacist, ethnocentric imperialists (Said, 2019, p. 207). Is such Orientalism not also racist? I believe you have addressed this in *Becoming a white antiracist* by using the term "BIPOC (Black, indigenous, and people of colour)" (Brookfield & Hess, 2021, p. xi) and of course, Orientalism is once again associated with white perpetrators.

Would you like to add to the definitory approaches on racism and white supremacy (and race) cited above? "Do you believe that racism is endemic and permanent" (Sheared et al., 2010, p. 23) in the U.S. as well as other nations and societies? Has it always existed, and can we never get rid of it? Is it possible to be in the apparently contradictory state of being "antiracist while also being in thrall of white supremacy" (Brookfield & Hess, 2021, p. 22)? Do you think that racism also exists in societies where Whites are not in the majority or dominant? Is it possible for racism to exist when Whites are not involved? Or can there be no racism without Whites? In other words: Is non-white racism a possibility? Is racism dependent on the amount of melanin of the victim and perpetrator? Sorry for the long question.

**SB:** That's okay. The first thing that I would say is that racism and white supremacy are two different things. But in the USA, in Western Europe and in the Northern Hemisphere, they are very much often conjoined together. But racism, as in the *Handbook of race and adult education* that you quoted back in 2010, is where one particular racial group entrenches its power and authority over other racial groups, either by overt violence and exclusion or by covert ideological manipulation – which is where the ideology of white supremacy comes in. But racism as a system of exclusion does not have to be practiced by white people.

*Racism occurs when one particular racial group entrenches its power and authority over other racial groups, either by overt violence and exclusion or by covert ideological manipulation... Racism is a system of exclusion and does not have to be practiced by white people.*

So, this speaks to several of your questions. Racism is this systemic, embedded exclusion, and that can be practiced by people of any skin color over another group that has a different skin pigmentation. So, it can be practiced within all kinds of black, brown, indigenous or Asian communities. It's not inherently a white European thing. It's just that the history of the world and the more recent world in the 18th, 19th and 20th centuries, the history of European colonialism and imperialism, means that the most glaring examples of racism are associated with white imperialism. But racism itself is just a system of structural exclusion based on racial identity, so that could happen potentially with any group at all. White supremacy is the kind of ideological justification of racism as it's practiced in the U.S. – I guess, I could say in Germany and certainly in the UK as well.

It's important to see those two terms as different. They're certainly conjoined and intersect in certain contexts, like in the United States. In the history of the world, basically, the one thing that you can guarantee is that at any point, some kind of genocide is being carried out by one group against another, based either on racial or ethnic identity or on other matters, possibly having to do with religious or tribal affiliation. So, this system of structural violence and exclusion is pretty much everywhere. Sometimes it shows up as being based around race, sometimes it shows up as being based around other identities, other kinds of characteristics. But then, when we look specifically at the world that I'm in, which is the United States, the most overwhelming form of racism is associated with whiteness and justified by white supremacy. That's the ideology that makes racism okay to a lot of people. So, you grow up with notions of whiteness that leadership looks like and being taught that the important historical characters who shaped the nation are the founding fathers of the United States. You see white supremacy there, and you see patriarchy as well, very much those two dominant ideologies. But if you look around the world, you have ethnic genocide.

*In the history of the world, the one thing that you can guarantee is that at any point, some kind of genocide is being carried out by one group against another, based either on racial or ethnic identity or on other matters, maybe having to do with religious or tribal affiliation. This system of structural violence and exclusion is pretty much everywhere.*

If you think of Bosnia or Rwanda, those are instances of ethnic genocide. You can see this in China and all around the world constantly. In my own life, I think of Northern Ireland, which was based on a religious divide, a desire to cleanse Northern Ireland of Catholics, make them leave and cross the border into Ireland itself, and to keep Northern Ireland part of the UK. I grew up in a system in which not only there was white supremacy, but also there was this ethnic stereotyping of Irish as less intelligent, as people who drank all the time, are constantly drunk, are constantly trying to avoid work. So, they were lazy. They were unintelligent. There were a lot of jokes about the lack of intelligence shown by Irish people. There were signs, when you went looking for places to rent in London, as I did in the early 1970s, saying “No Irish, no blacks, no dogs” [laughs].



Figure 6. A sign reading “No Irish, No Blacks, No Dogs”. Source: Draper (2015).

At this time, you could do that without legal penalty. And then when you look at communities of color, I’ve often heard my colleagues and friends of African descent tell me that there’s clearly white supremacy within those communities: so that the lighter-skinned you are, as broadly speaking a person of African descent, the more desirable your views are, the more intelligent or, the more befitted you are for leadership. So, if you have a child that’s born very dark and another child is born very white, you’re usually happy about the very white child because there is this sense that the barriers will be slightly less to them. And they’re considered more attractive and more beautiful. That colorism also comes from white supremacy.

### *Colorism comes from white supremacy.*

My particular focus on white supremacy is because for the last 40 years, I’ve lived and worked in the USA, and that’s very clearly the predominant form of racism that exists. It exists to justify a system of structural exclusion that you cannot avoid witnessing in the United States.

But if I were in a different context, there would be other forms of structural exclusion based on racial identity or within the same racial group based on ethnic identity. This is just part and parcel of what it means when humans organize themselves in groups and live together. This seems to be a constant feature of history. In terms of that question, ‘Do I regard this stuff as endemic and permanent?’ Yes, I do. It’s very clearly the case in the parts of the world that I know,

which are Western Europe and the United States. I would say broadly that white Commonwealth nations and the other nations that have been colonized in the 17th, 18th, and 19th centuries, all have a lot of white supremacy still embedded in their culture. So, I do regard it as permanent and endemic, and of course, that’s part of a critical theory worldview. That’s the first thing you look for as you think about your world. You look for structural inequity and how that’s made to seem normal. So, it’s not a big surprise to me or a big leap to see the world in that way.

*Racism is endemic and permanent... I do think it’s always existed. Can we ever get rid of it? As with many things, even though historical and empirical evidence disproves the possibility, you have to act as if the possibility is true and realizable.*

I do think it’s always existed. Can we ever get rid of it? As with many things, even though historical and empirical evidence disproves the possibility, you have to act as if the possibility is true and realizable. So, can we ever get rid of it? I don’t know. But we have to try; we have to believe that that is possible and that it can constantly be pushed back and diminished because I’m essentially, I guess, a modernist. I have always believed that progress is possible, but I also realize that any progress will be resisted and will be pushed back against very severely.

Is it possible to be in the apparently contradictory state of being antiracist while also being enmeshed in white supremacy? Yes, because I am. In *Becoming a white antiracist*, that’s what Mary Hess and I were arguing: The two of us will never lose the fact that we’ve been soaked and indoctrinated and learned this white supremacy at a deep level from an early age. But you can recognize that that’s the case and think about how you’re going to push back against it, how you’re going to limit its influence in your own actions, but particularly how you can help organizations and institutions and communities name it as an important reality.

*Is it possible to be in the apparently contradictory state of being antiracist while also being enthralled by white supremacy? Yes, because I am.*

I think about how we’re going to try and dismantle it in the best way that we can. Do I think racism exists in societies where non-Whites are not in the majority or dominance? I think that structural exclusion exists in those societies. Sometimes, yes, that is based on race; sometimes, it’s based in racially homogenous societies around class or gender or ethnicity, those other dimensions. Is it possible for racism to exist where Whites are not involved? Yes, it really depends on which racial group is dominant. That is why, in the United States, when I talk about racism, I always have the notion of white supremacy there. So, it’s racism and white supremacy: racism – the structure of exclusion – and white supremacy – the ideological justification for that structure to be in place in this country.

I always try to remember and tell people that you also have to understand that racism is usually linked to some kind of political or economic project. In order to justify treating one group in an inhumane way, putting them through horrendous working conditions, and exploiting their labor, you need to view them as less than human; you need to view them as expendable animals. Slavery, in many ways, built the economy in the United States; there was a need for cheap labor – not just cheap labor, but free labor. For white Christians to justify that, you have to carry this idea in your head that ‘Well, when the Bible talks about people, they’re not really referring to, let’s say, black and brown or indigenous peoples because those aren’t really people. They’re subhuman. They’re animals. That’s why it’s fine for us to exploit them because we’re not contravening God’s will or we’re not contradicting the teachings of Christ.’

Is non-white racism a possibility? Yes, non-white racism is definitely a possibility, depending on the context and geography. Is racism dependent on the amount of melanin? In the United States, that’s how it’s constructed. Genetically, across the world, there’s hardly any difference between humans – at least, that’s where we are in our understanding of genetics right now – the differences are extremely minor. But racism is this construct that people erect to justify treating another group in the ways we described and exploiting them for their own benefit. So, racism is not dependent on the amount of melanin. But racism in the United States, because it’s associated with white supremacy, is dependent on the amount of melanin in the victim and perpetrators. That’s how we decide that another group can be treated in an inhumane way: because they do not look white.

*Genetically, across the world, there’s hardly any difference between humans – at least, that’s where we are in our understanding of genetics right now – the differences are extremely minor.*

JR: I had some nagging doubts when reading *Teaching race* (Brookfield & Associates, 2018) and *Becoming a white antiracist* (Brookfield & Hess, 2021). Having this conversation with you today is really so clarifying. We found your discussion of the various types of racist violence (structural, cultural and direct) most insightful. Citing Galtung (1969, pp. 171, 191), the structural violence of racism manifests itself as “unequal power and consequently... unequal life chances”, while the cultural violence of racism refers to cultural aspects (assigning individual characteristics such as ‘inferior’, ‘lazy’, ‘stupid’, or ‘inherently violent’) “that can be used to justify or legitimate direct or structural violence”. Structural and cultural violence are then “used to justify direct violence, as housing is destroyed in gentrification, calls for justice are repressed as riots and unarmed people of colour are disproportionately killed by police” (Klein, 2018, p. 105). Another indication of systemic racism is the mass incarceration of black people. Is violence in communities of colour a symptom of (rather than a cause for) poverty and a reaction to the different types of racist violence outlined above?

SB: I’m going to speak about the USA, first of all, and I would say that the celebration of violence in this culture is very clear to me as someone who didn’t live here. I moved here when I was 33. For the last 40 years, I have lived here. It is very clear that the deification of guns, the Second Amendment (the right to bear arms), and the belief that everybody is almost like a frontier settler who needs to defend themselves against the ‘savages’ (that was the word that was used to refer to the tribal nations, the indigenous people in the U.S.) is how people typically grow up. That glorification of violence is not limited to any community at all. It is very much an American value.

So, I think we have to understand this in the context of the whole American cultural celebration of violence. Growing up, I’m sure, you in Germany as I in the United Kingdom, we saw a lot of Westerns where the pure white settlers were shooting and killing Indians just indiscriminately from horses as they were attacking the wagon train or from a farm homestead. So, it’s very much a part of the American psyche. I wanted to say that first of all.

It’s a very dangerous thing when white people make generalizations about, in particular, black or brown culture. There’s that part of how white supremacy perpetuates itself because there are a lot of generalizations and stereotypes that whites invoke about communities that they’ve never visited. Even the police overwhelmingly live in communities other than those that they’re responsible for policing, and they regard these communities as homes for intentionally violent people, full of criminals, lacking intelligence, and so on and so forth. But I can speak about the question as a white person because I have experience of how white supremacy is learned, and I’m sure I’ve enacted multiple times the perpetuation of these kinds of stereotypes. It’s easy for me to fall into them.

That’s why it’s so important for me to have constant contact and work with and try to live with people of color so that I get this counter-narrative. So, if you look at media in the United States, they have made an enormous amount of money through rap, hip hop and gangsta rap, which celebrates gang life and people killing each other, depending on the particular gang that they’re in. So, that whole genre of gangsta rap celebrates the particularly black-on-black, black-on-brown, brown-on-black, brown-on-brown kinds of violence.

What you don’t get are other images of black life that you encounter when you actually talk and live and work with black or brown people: essentially, their life being focused on the church is a major part of life and of the community; it’s focused on deep community, it’s focused on incorporating spirituality into their lives. There’s this very strong sense of fierce collectivism, ‘we’re going to get through this together’, a strong emphasis on loving relationships, on family, on food.

All of those things are kind of on the periphery of white consciousness of black life. Because they counter the narrative that white supremacy has taught us, which is essentially that folks of color are constantly on the verge of exploding into some kind of violent criminal activity.



Figure 7: Cover art for N.W.A.'s album *Straight Outta Compton*, the first blockbuster gangsta rap album released in 1988. Album Cover art and design by Helane Freeman. The cover art copyright is believed to belong to the label, Ruthless Records and Priority Records, or the graphic artist(s). Fair use.

As I said, I don't want to talk about the reality in those communities. But I can talk about the reality of how white supremacy has structured a white view of what goes on in those communities and how, in terms of my limited experience, the reality in those communities is much fuller and richer. Certainly, the communities are worried about violence, as any community would be. But the reality has much more to do with collective pride, with relationships, with family, with holding together in the face of a sustained onslaught, with staying very vital and alive through music and song and dance and food and the incredibly vibrant artistic forms and representations that you find in black and brown communities. That stuff tends not to be featured because it doesn't fit the white supremacist narrative.

### The silencing of critical race theory

ST: We now propose to turn our attention to Critical Race Theory (CRT). The term *intersectionality* refers to the interconnected nature of social categorisations such as race, class, gender, sexual orientation, ability and age as they apply to a given individual (or group), thus creating overlapping and interdependent systems of discrimination and disadvantage. Identities are complex and plural. Mike Klein argues against "black and white (pun intended)" categories that prevail in popular culture. He states that race is socially constructed and complicated by categories such as "ethnicity, nationality and hybridity" (Klein, 2018, p. 101).

Is there a need for an intersectional perspective (that addresses racism in addition to classism, sexism, ageism and ableism)? Would you agree that you have advanced

critical theory by incorporating Cornel West, Bell Hooks and Angela Davis in your own magisterial interpretation of critical theory (Brookfield, 2005)? What are, in your view, the main aspects of CRT? For the last couple of years, conservative U.S. lawmakers have sought to ban or restrict CRT from primary and secondary schools (e.g. in Idaho, Iowa, Oklahoma, Tennessee and Texas). Why do they want to silence discussions of racism, equality, social justice, and the history of race in the classroom? Again, apologies for the barrage of questions.

SB: Yeah, but they're all good ones and things I think about a lot. In fact, I recently wrote a foreword to a book, which is a conversation around how we integrate a class analysis of inequality with a racial analysis (Brookfield, 2023). And do they need to be separated at some point? Or should they always be an intersectional analysis? This is in my head right now.

Is there a need for an intersectional perspective addressing racism in addition to these other isms? Yes, of course, I don't see how anyone could not see that an intersectional perspective is important. However, I say that with a qualification, which is that in the United States, where obviously I'm located, people are generally more comfortable talking about even sexism – and the patriarchy that justifies it – and classism, ableism, and ageism. Those are isms, generally, that it's easier to talk about. At least, this has been my experience.

Maybe I'm just speaking in an anecdotal way that others can't support. But in my own work, I have noticed that the hardest thing to get people to do is to focus on race as a category of analysis. All the other isms that justify structural exclusion are much easier to get the conversation around. So, when you do an intersection analysis with a group, what you have to watch out for is race being lost in that. So, that means that as an educator, I have to foreground race. I believe I do anyway; I've constantly drawn attention to this. I sometimes say, 'Well, these other isms are important, but we need to focus on race right now because that's the thing that keeps getting lost'. As long as people are unwilling to talk about it and don't know how to talk about it, it's very difficult to know how to address the ism associated with racial identity. So, I see the complete validity of intersectional analysis, and I agree with it, but I also know that talking about race is the hardest thing for many to do. So, I need to keep focusing on that particular ism, because it's the one that will get lost if I don't do that.

*I see the complete validity of intersectional analysis, and I agree with it, but I also know that talking about race is the hardest thing for many to do.*

In *The power of critical theory* (Brookfield, 2005), I wanted to do an intersectional analysis. I deliberately have a chapter in that book on sexism and patriarchy. I have a chapter that you cite on racism and white supremacy. I wanted to show how that critical theory tradition was still incredibly accurate, helpful and relevant for the late 20th and early 21st centuries. So, that's why I had those chapters in that book. Again, I feel that when we look at critical theory, structural inequity and

so on, it is easy to lose race in that analysis. I didn't want that to happen with my own work.

But I do sometimes get criticized for neglecting all the other things by what many people see as the consistent focus on race that I have had in the last few years. But again, that's deliberately a strategic decision. My reason for doing that as a white person is to give a little example of how Whites can engage in a critical theory-influenced analysis which focuses exclusively on race.

What are the main aspects of critical race theory for me and their tie to critical theory? First of all, there's this view of racism as endemic. Critical theory views exclusionary isms as endemic, and critical race theory views racism as endemic. So, that's clearly an obvious axiom for me. Critical race theory has really advanced the idea of counter-narratives and using personal experience and personal testimony as an educational tool to get people to look at race. So, I am a strong proponent of using narrative. I find it less than ideal to start off with statistics of the school-to-prison pipeline, disproportionate access to education, the way that COVID has disproportionately affected communities around race or how toxic waste dumps typically tend to be sighted in communities of color – you can talk about all those things and quote statistics. But it's not as powerful a tool, educationally speaking, as hearing a compelling story, hearing an individual talk about how that affected their life. That's one reason why I use a lot of digital stories, things I find online as beginning points of access into looking at race.

Then, after a while, once you've got people's attention, you can step back, and then bring in the statistics and the theoretical stuff. So, that's the second part of CRT, which has really been influential on me. The principle of interest convergence is a very important insight: this idea that massive, permanent structural change will only come about when Whites see it as being in their own self-interest. It speaks to the nature of a multiracial movement or a multiracial alliance. It speaks to the point that Whites have to be involved in this work in different ways, but they are absolutely necessary and important to it. It also speaks to what I was talking about earlier when I was trying to argue that it's in our own self-interest to get rid of these toxic ideas because if we believe them, we're engaging in all kinds of incredible intellectual gymnastics to convince ourselves that Whites really are the inherently superior group [all laugh]. Or we're living in constant fear of anyone who doesn't look like us.

*The principle of interest convergence is a very important insight: this idea that massive, permanent structural change will only come about when Whites see it as being in their own self-interest. It speaks to the nature of a multiracial movement or a multiracial alliance.*

That principle of interest convergence has been very important to me. Finally, CRT's emphasis on intersectionality is very important because that has always been a consistent

part of any CRT analysis I've read. Even though they're talking about race, they're also saying 'race is not the whole reality; there are other systems of exclusion based on different positionalities and identities that are in play in the United States'.

When you ask about attempts to ban critical race theory, it's quite incredible to me that we have executive orders banning the mention of critical race theory, as we did under the last President Donald Trump. You could not refer to critical race theory and certainly not teach it. Not even name it and mention it in any kind of federal training! So, there we have a direct example of state ideological control in play: very clear, naked, nothing covert about it, it really laid out the battle lines.

That was partly an attempt to play to a very conservative, evangelical, right-wing base. But also, it's symptomatic of the real fear that a lot of those in the elites in power feel, that now this race stuff is getting out of hand. In the past, you could have demonstrations. You can convince people that legislation has taken care of the problem when it really hasn't; it's just reconfigured the ways in which this permanent endemic racism is enacted.



Figure 8: Tributes and mural outside Cup Foods, where Floyd was murdered. Photo by. Vasanth Rajkumar. CC BY-SA 4.0.

There's this sense that President Obama's election, first of all, was a real challenge, but his policies were mostly centrist. But there was such alarm triggered by having a black president that it ushered in this enormous right-wing wave of fervor and hatred of anything black and a desire to dismantle everything that Barack Obama had done. Trump's and the Republican Party's legislative agenda was an attempt to turn the clock back on everything that Obama had managed to do during his eight years. You had all that, and then you have the film of George Floyd dying in May 2020, minutes and minutes passing, while you see a man fall into unconsciousness, complaining that he can't breathe. You see that psychopathic stare from the police in their attempts to hold people back and not render any help. So, you have this wave of outrage and the possibility of mass mobilization across racial lines because there are a lot of white members showing up to Black Lives Matter protests and participating in those in different ways.

The reason why CRT is being banned is that now, there's a real sense of fear on the part of those rich white elites who disproportionately control resources and economic life. When you look at where these attempts to ban CRT come from and who they're funded by, you trace the money. It comes from white billionaires behind enormous companies in the United States who fund all this stuff. If you've been reading about the fact that Canada had this truck convoy that just paralysed Ottawa and trade in the eastern part of Canada, you trace the funding. It's from U.S. corporations, these rich white elites who own enormous resources. It's a sense of a realistic threat to the status quo that is behind these attempts to quash that threat and make sure the status quo is not really challenged at any fundamental level. And the way that it's happened is the old ideological trick to conflate a certain point of view with being un-American, anti-American or unpatriotic.

Now schools are not allowed to teach perspectives that mention white racism, white supremacy or the exploitation of other races by Whites. Because that's seen as unpatriotic and countering the dominant narrative of the United States, which is: 'We are an evolving democracy in which a meritocratic system operates, and everybody has the possibility to flourish'. So, as critical race theory would say, that's the official story, the official narrative, but the counter-narrative is becoming much more widespread, especially amongst the young. I guess rich white elites say, 'We got to do something about it. Desperate times call for desperate measures. So, we'll ban this perspective as part of public education in the United States, and we'll ban this perspective in terms of any federal training'.

There were times in the last days of the Trump era when I was doing my work around anti-racism, and I was having to work within the confines, certainly in any federal stuff. But even not in the federal government, the influence of that order not to use CRT got people very spooked in all kinds of institutions, particularly public institutions. I would get lots of questions: 'When are you going to reference critical race theory?'

I take it mostly as a sign that rich white elites realize we're at a potential turning point here in this country, and so the divisions and the battle lines are being drawn ever more clearly. That's not always a bad thing because once those lines are clear, you have to choose a side; you have to say which side you're on and which side you're going to support and work for. You can't go through your life thinking: 'Well, we're making progress. The civil rights era has brought us a long way and things are not as bad anymore as they were 40, 50, 60 or 70 years ago'. As a white person, it's easy to carry that narrative in your head. I don't think you can carry that narrative anymore. So, there's a clear cultural war going on here in the United States, and there always has been; that's part of how an elite maintains its power by manipulating the ideology that's prevalent in a culture, but usually, it's not done so overtly as it is now.

## Is cosmopolitanism a viable alternative to racism?

JR: Your co-author, Mary Hess (2018), discusses cosmopolitanism in *Teaching race*: one "can become, indeed should aspire to be, a citizen of the world, able to embrace local ties and commitments, but also to extend well beyond them, engaging a wider human community, even across divides of seemingly irreconcilable differences" (Avila & Pandya, cited in Hess, 2018, p. 270). Also, Kwame Anthony Appiah (2006) articulates a cosmopolitan community where individuals from varying physical or economic locations enter relationships of mutual respect despite their differing political or religious beliefs. Is cosmopolitanism a viable alternative to racism inasmuch as it assumes that all human beings are members of a single community? Or is cosmopolitanism yet another privileged perspective that is class-based?

SB: Another very interesting question! I would just start off by repeating my assertion that in terms of genetics and biology, we all are members of a single community. There is so much more genetically that unites us than divides us across the world. So, again, notions of racial, ethnic, and other divisions are entirely human constructions that have nothing to do with biology. I haven't really used the term *cosmopolitanism* in my own work, but one of the things that I learned from reading more deeply and talking to others about the Afrocentric perspective is that that perspective, even though it is a culturally-based one, recognizes and honors the validity of other culturally based paradigms. An Afrocentric perspective, briefly defined, is one grounded in African cultural values of collectivism in particular, rather than in a Eurocentric perspective where individual independence, solo, critical thinking, being the captain of your own ship, of your own soul, constructing your own life as an individual, that's the paradigm there. The Afrocentric paradigm and also other paradigms – indigenous, tribal, and even working-class perspectives – are much more collectively-based.

*In terms of genetics and biology, we all are members of a single community; there is so much more genetics that unites us than divides us across the world. Notions of racial, ethnic, and other divisions are entirely human constructions that have nothing to do with biology.*

One thing the Afrocentric paradigm emphasizes is that it is just one way of looking at the world alongside many others and that there are multiple ways of understanding the world. None has any inherent superiority or validity. I really liked that part of Afrocentrism. It did not proclaim that it wanted to replace European or Eurocentric epistemology; it just said, 'We have a different set of interests and worldviews that are counter to what tends to get privileged in Eurocentrism'. Politically and ecologically, we're interdependent, whether we like it or not.

You can't really separate interests out in an independent way because we are interdependent with each other. We're interdependent with the world, with a natural world, with Mother Earth, with the ecosystem. So, we're seeing the

effects of living in a Eurocentric way and the influence of positivist epistemology, which separates humans from the land, separates humans from nature, and assumes that nature is there to be controlled and manipulated for human advantage. The consequences of that are now really coming home to roost and threatening the continued existence of everybody in the world. So, if anything should prove the truth of independence, you would think it would be the damage to the ecosystem that's happening.

My own preference is to focus on collectivism; I tend to use that word a lot rather than something like cosmopolitanism. In the US, collectivism is not a privileged perspective. Certainly, we talk about being good neighbors, and there is this myth that in crisis, everybody will help each other, and we'll get through this. But in reality, it is one class, one race, one ethnicity pitted against the other, sometimes very deliberately, as part of how white supremacy keeps itself unchallenged.

Collectivism is often associated as socialistic or even communistic, as un-American, anti-American. So, in the United States, the word *communitarian* is used a lot. There's been a lot of stuff in the '90s around communitarian philosophy and the need to re-animate communitarianism within the United States as a counter to this fervent, rabid individualism that is so much baked into the American cultural pie. In terms of empirical reality, we are interdependent, and there isn't that much that divides us biologically speaking; we are a single-world community. The things that divide us are humanly constructed, and therefore, if something is humanly constructed, logically, it can be humanly deconstructed and reconstructed. This goes back to your earlier question: 'Is racism endemic?' Well, yes, it is, but it's been humanly constructed. Therefore, ending it is logically possible. If something has been constructed, it can be deconstructed and reconstructed.

## Racism in higher education

**ST:** In *Teaching race* (Brookfield, 2018, p. 2), you state that racism is glaringly obvious in several aspects of the educational sphere: in admissions policies, disciplinary guidelines, curricula, hiring practices, attrition rates for faculty and students of colour, and the composition of boards of trustees. At the same time, Marcuse's (1969) concept of repressive tolerance is also applicable to race in higher education:

"By allowing a limited amount of protest that is carefully managed, a societal pressure valve is created to release into thin air the real change. Diversity days, Black History Month..., colleges and universities featuring photos of Black, Brown, or Asian students on their publicity materials (when such students comprise only a small minority of actual students)... - all these can be seen as examples of repressive tolerance" (Brookfield & Holst, 2011, p. 28).

Could you please elaborate on your insights and observations on racism in the educational sector?

**SB:** I speak again from personal experience here. In a lot of different institutional contexts, there's a very predictable course of events that I've witnessed over the years: an institution – it could be a college, university, school, corporation, government agency, religious organization, military or any kind of organization – is accused by a particular group of systemic racism or hate crimes are highlighted that have happened on site. This becomes public. Now the institution feels 'We have to do something to demonstrate to the world that we're taking race seriously, we're not a racist organization'. There are these very predictable things that happen.

The first is the public relations work – you've already quoted John Holst and me referencing this – where you work on your website, on your Instagram or Twitter accounts, on the brochures, your admissions pages, and your alumni magazines, which in the U.S. is a very important source of raising money. All these things now feature a rainbow coalition of different racial identities that make it look as though your organization has nicely balanced, equal proportions of black, brown, white, indigenous, and Asian faces. So, that's the first thing that happens. It's a total fiction. I've spoken to many individual students or faculty who are sick of having their faces plastered and being highlighted and featured because they recognize this as a con job. This is a public relations manipulation and nothing more. So, that's very much an example of racism that looks like it's antiracist. It is actually racism in the sense that it's perpetuating the system and keeping a fundamental challenge to the system at bay.

A lot of institutions now – I don't really know of any institutions that don't – proclaim 'We are about diversity, equity and inclusion'. Sometimes, they'll go as far as saying: 'We are an antiracist institution, and here are all the workshops that we've run to justify that claim'. So, you put on a lot of anti-bias or microaggressions workshops on different topics around diversity and inclusion. And it looks like you're really taking this seriously because 'look at all these workshops'. But again, nothing fundamental is being changed in regard to admissions, funding priorities, how student work is being assessed, and what behaviors are taken into account when promotion, retention, and tenure decisions are taking place.

Another thing that happens – again, which looks antiracist, but in fact, is a way of deflecting antiracism – is that you appoint some people of color to an influential position. So, now you can say, 'Look, we have a diversity office run by this particular individual who is a person of color, so clearly, we take this seriously'. Or you drop a body of color into different units across the institution, and you highlight that 'Well, now we have a person of color in mathematics or in biology or whatever the discipline is. You see, we're really taking racism seriously, and we're trying to be antiracist'. The problem is that you drop those people into this sea of whiteness with no support for them. They're the only black, brown or indigenous person in the department. They're the ones who have always volunteered to serve on diversity committees or task forces to represent your department. So,

they have this whole load of extra work to do simply because of their racial identity. But you don't give them the support. So, after a couple of years, they burn out, and they leave. I've seen this happen over and over again. Because essentially, when you appoint this person, you're saying, 'Yes, we want you to teach mathematics or biology, but we also want you to take on a second full-time job of being someone who will educate the rest of us Whites about what racism is and how we can push back against it. But we're not going to name that as a job responsibility, and we're not going to pay you anything. So basically, you have two jobs, but you don't get paid for one'. I've had a lot of colleagues of color talk about that frustration to me and how they're always the ones who are asked to serve on diversity committees.

Another thing is that as you're setting up diversity initiatives, you frame them. This may be reiterating the point I've just made, but you frame them as the responsibility of people of color. Most diversity offices are headed by a person of color; you'll rarely see a white person in charge. In one sense, that's completely understandable. Because a white person doesn't have the experience of being on the other end of racism. On the other hand, as we were talking about earlier, they do have the experience of how you enact, learn and reproduce white supremacy as normal behavior in your life. I'm always advocating that a multiracial team should run a diversity office, and one of the team members should be white. Because in the mix, we need to understand the continuing presence of unexamined white supremacy, that continuing refusal to look at what it means to be white whom institutions will have – particularly predominantly white institutions. You've got to break with that. The only way you can break with this is to have a multiracial conversation involving Whites as well as people of color about the way that white supremacy keeps reproducing itself.

As long as you keep white people out of that picture, you can successfully continue the idea that race really is a problem for people of color. But in fact, people from James Baldwin onward have made the point: 'No, the problem of race is the problem of white supremacy, unacknowledged white identity'. The problem of racism is a white problem. It's something that I had a lot of experience with and feel passionate about.

Another thing that I see happening is: 'We'll do diversity workshops'. Most people are very happy with the terms *diversity* or *inclusion*. 'We'll focus on diversity and inclusion, we'll celebrate all the different identities that we have and the different parts of the world. People on campus come from the different experiences they bring; we will enrich each other by learning about our different cultural identities and histories'.

Absent from that framing is racism or white supremacy. You can do a whole diversity and inclusion initiative – it's harder to do it around equity – you can do that by celebrating human difference and completely eliding any reference to racism. That way, it looks like you're addressing racism, but you're really just again avoiding tackling the issue in any meaningful way.

*People from James Baldwin onward have made the point that the problem of race is the problem of white supremacy, unacknowledged white identity. The problem of racism is a white problem.*

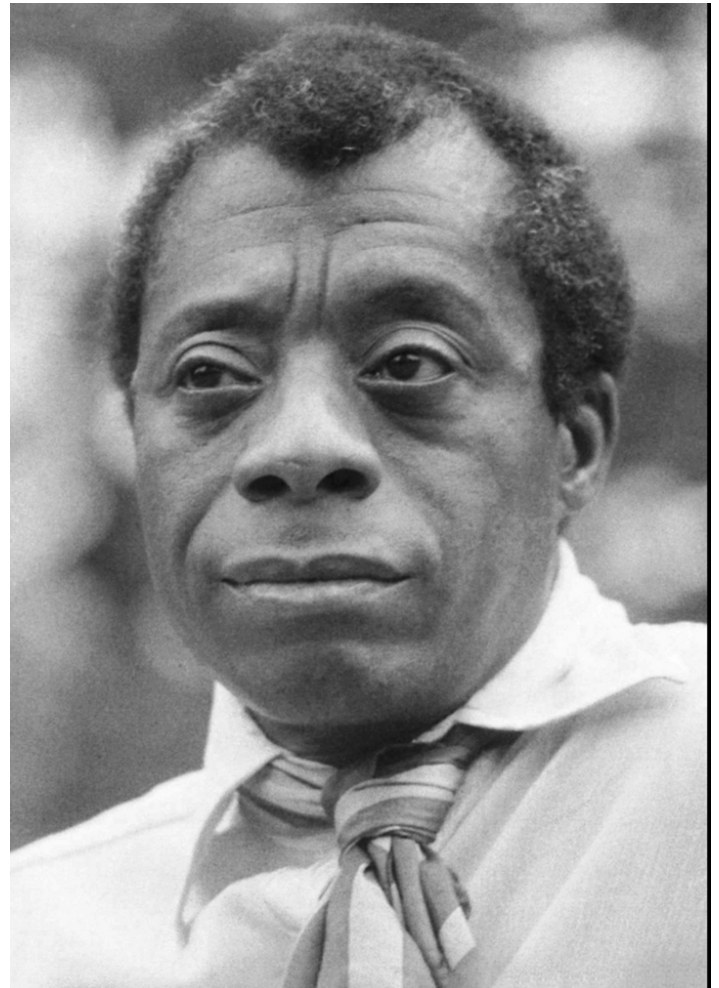


Figure 9: James Baldwin in 1969. Photo by Allan Warren. CC BY-SA 3.0. Baldwin (1924 – 1987) was an American writer and activist.

In higher education, Eurocentric epistemology is almost completely unchallenged. There are many things about Eurocentric epistemology that I like and that I think work to people's benefit. But it's not the only way of establishing knowledge and coming to the truth. For example, it's very difficult to do collaborative assessed work in higher ed, to do a group project. Yes, it will be allowed occasionally, but with transcripts of student achievement or the way that we structure dissertations if you go into graduate-level work, this is all done very much within a solo individual, scholarly paradigm. This extends into faculty work as well, where co-authored publications are not viewed as valuable as solo-authored publications; where if you want to do a team-teaching course, for budgetary reasons, you only get half a load or a third of the load, depending on how many people you're teaching with. Disciplines are silo'd, it's hard and a lot of work to get interdisciplinary courses off the ground.



Yet, the extra work associated with those is not rewarded. In fact, if an interdisciplinary course is team-taught, then you're penalized for that effort. The whole privileging of text, of the written word, of solo scholarship, of assessments being done individually, silo'd disciplines, team teaching being hard, sometimes arguing 'Well, in our discipline, we are engaged in a non-political search for truth'. The denial of any political project behind which disciplines get funded, the fact that STEM disciplines are disproportionately funded, is clearly a response to the desire by, again, rich white elites to have people schooled in these particular skill sets. It's a political choice, but it's presented as apolitical so that whole Eurocentric epistemology is pretty much unchallenged. It's very hard to get, for example – I've tried this myself in my own world – a collaboratively-authored doctoral thesis.

Then, just to finish up this long answer: the other thing that happens is that those with the real levers of power, who are the trustees or the governors that appoint the President, determine the strategic direction of a particular higher education institution in the United States. Their workings are completely secret to the whole community. It almost exists as if they don't exist. People think that the President of the institution is the one setting the policy, but they're not. The president serves at the appointment and pleasure of the Board of Trustees or the Board of Governors. They operate behind this cloak of secrecy. If you really want to get antiracism addressed, they have to come out and be part of a public conversation around it. I hardly ever see that happening.

You can do all these diversity and equity initiatives without changing Eurocentric epistemology or the power of the trustees or the Board of Governors, who are responsible for setting the direction and tone of the institution. You can do it and say you're anti-racist. But what you get rewarded for in terms of getting promoted, in terms of how you're appointed initially to a position, especially a faculty position, it's just the same old stuff: Eurocentric-epistemologically-determined, scholarly accomplishments such as solo authorship. You get tenure by racking up the number of books and articles that you've written on your own, at least in a Research One institution [so-called R1 or Doctoral Universities in the U.S., characterized by 'very high research activity'] that in itself is positioned as the most admired and rigorous higher education institution.

If you're a teaching institution in the United States, you're usually regarded as second, third or fourth tier. The ones that US News and World Report feature as the most prominent tend to be Ivy League, Big Ten universities and Research One universities, where you can be a terrible teacher. If you publish in refereed journals, then that's what brings you the status.

Anti-racist pedagogy is nested in a system, an institution, and an organizational culture. You can do your own individual stuff in a course or in a particular class, but unless you address the way the culture influences and frames what's going on and the structures that are in place, and the policies and the reward systems, unless you do that you're not really changing anything. It's kind of performative.

## How can we teach about race?

JR: While engaged in teaching race, "white guilt is not the desired educational outcome" (Smith, 2018, p. 187). Klein (2018) emphasizes that the point of such a critical race pedagogy is "not to assign blame or wallow in guilt, but to critically assess normative assumptions and to free ourselves from racist social constructions so we can pursue education as the practice of freedom" (p. 89).

How can we come up with a series of sequenced stages to bring students into discussions where their identities come into question, particularly in predominantly White institutions? How can White teachers address the topic of race in ways that don't re-centre their power? To "paraphrase Marx's 11th thesis on Feuerbach", how can we not just understand how racism works but "seek to change it" (Brookfield & Hess, 2021, p. 40)?

SB: I'm going to break the question down and answer the three subsections in order. The first one is about bringing students into discussions and the sequence of stages: I have nine or ten things that I typically do in often the same sort of sequence, so I'll go through those. Of course, all of these change according to the specific context you find yourself in. But I think the first step, when you're getting ready to work with students, even before you meet them, is to reframe what counts as success, to try and lose your desire or your expectation to do it correctly or perfectly. This all goes back to that line I like to quote: 'There are two ways to do antiracist work: imperfectly or not at all'.

*There are two ways to do antiracist work: imperfectly or not at all.*

One of the things that demoralizes those of us going into this work early on is that we have this image of what a good workshop, course, training or meeting looks like when it's trying to incorporate issues of race into the discussion. But frequently, when we enact that practice, the reality is so far from our imagining that it gets very demoralizing because you have a false sense of how it looks to do the work well. This usually means that you can see a smooth upward trajectory in terms of people's understanding, that they get better at recognizing their assumptions about being aware when they're engaging in micro-aggressions. They start to raise race as an issue in class and outside, and things stay on an even keel emotionally. There's no awkward silence, and everybody participates roughly equally, offering the same amount of contributions.

So, that will be the first thing to get rid of: that set of expectations because it won't be like that. There will be a lot of emotions and a lot of strong feelings expressed; there will be long, awkward silences when people don't know what to say or whether to say what's on their minds. There will be expressions of frustration, anger, and sadness. Just remember that as you're going into this, it will not match some earlier experiences you've had teaching content where race is not involved.

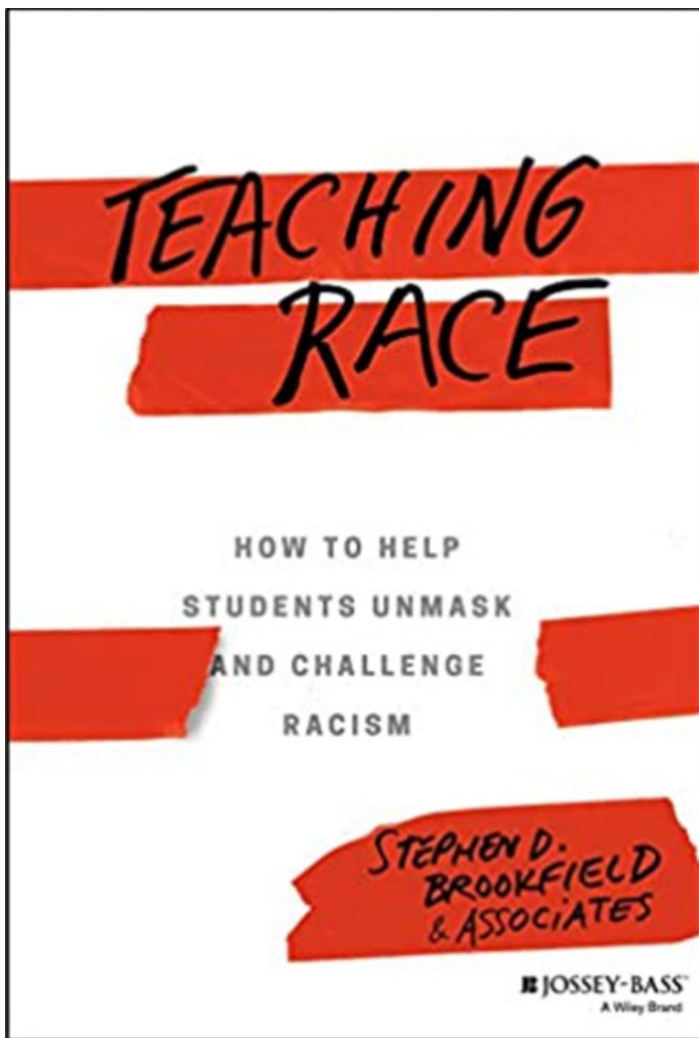


Figure 10: Book cover of *Teaching race* (Brookfield & Associates, 2018).

Then, when you've done that mental preparatory work and you meet with students for the first time, you have to do a lot of self-disclosure about your own racial identity, the role which race has played in your life, the struggle that you've had in terms of taking race seriously and understanding that if you're a white person and you have a white identity that really signifies something to the world. The first encounters with students should have a lot of narrative self-disclosure; you can't ask anybody to talk about racial experiences or issues until you've first done it publicly, several times. Then, when you start to enact a curriculum, conduct training, or run your task force meetings, it's always better if this is done as a team of facilitators who can model racial crosstalk. So whenever possible, I like to give students some early exposure to myself with at least one other colleague who comes hopefully from a different racial identity than my own. If students can see the two of us modeling, talking around race, leaving plenty of silence and talking about how we still struggle with these issues all the time, I think that is a very helpful tone-setting piece of modeling that facilitators can do.

Then I would bring in as a fourth step, some digital narratives, some contemporary examples that you find on YouTube or some other social media site – TikTok, Instagram, whatever it is – and you can get people to view some personal testimony

about the nature of racism, and the effect that gives a central focus often at the beginning of a course or a workshop that people can begin to work from. So, instead of asking people to share their own narratives initially, I would probably be using a lot of digital narratives early on.

Next, I would constantly be taking the emotional temperature of how things were progressing: that ability to access students' experiences, to get accurate information on how they're experiencing, what's happening to them, the learning that's going on in groups, the way that they're interacting with each other, the way they're interacting with you, their reactions to the content. If you have regular information about how students are experiencing those things, then it just makes your choices much more grounded in reality than they otherwise would be.

Next, you have to introduce the concept of brave space – or if you don't use the brave space language, just alert students to the fact that when we engage in looking at race, particularly in a multiracial classroom, there will be all the things that I mentioned earlier: there will be long silences, there will be expressions of emotion, feelings, anger, sadness, and frustration. If students don't understand that stopping for a long time and nobody saying anything while they think it is quite normal – if they don't have that sense when they go into it, then they're constantly going to be feeling like, 'Well, the instructor doesn't know what they're doing, they've lost control, we're spiraling into the expression of emotions' and so on. You have to prepare students for the nature of racialized conversation; you can talk about potential ground rules and show examples. Again, I'd use a lot of digital examples of what racialized discussions will look like, which involve deep expressions of emotions and feelings.

The next thing you do is start considering when it might be good to split students up into racial affinity groups so that they have time to interact with others from their own racial identity. That stops a lot of the dangers we've already talked about when we have new multiracial groups talking about race. If it's people of color in one group, they can speak honestly about their own experiences of racism with each other in a way that they might be constrained about in a group comprising a lot of white students as well. In a mixed-racial identity group, one of the things that happens often is that white participants doubt the veracity of the expressions of racism that come from people of color. They're trying to talk them out of it and say they're too sensitive. 'That's really not what was going on' and so on. You don't have any of that if you're in a group with other people of color, plus you don't have white members trying to show you how they are allies or constantly asking you for information and advice and to teach them and to tell them what to do.

On the converse, in a white affinity group, if you don't have members of color, then the whole temptation to perform your wokeness for members of color is gone. You can really just talk about what whiteness means to each other in a more relaxed way than you would if it was a group which comprised multiple racial identities. I think the affinity group strategy is important to incorporate at various times. I don't mean that you keep people in those groups for the whole

class or the whole course. But I do think that a strategic use of how you group students into small group discussions around racial identities is important.

Then, when you get into classroom activities, I'm a very strong advocate of conversational protocols. Some of the ones that I use very frequently when I'm teaching in this area are Circle of Voices, Circular Response, the Chalk Talk approach, and an approach based on David Bohm's work, the theoretical physicist, that Steve Preskill and I in a book on discussion that we wrote call Bohmian Dialogue (see Brookfield & Preskill, 2012; 2016; Brookfield et al., 2024).

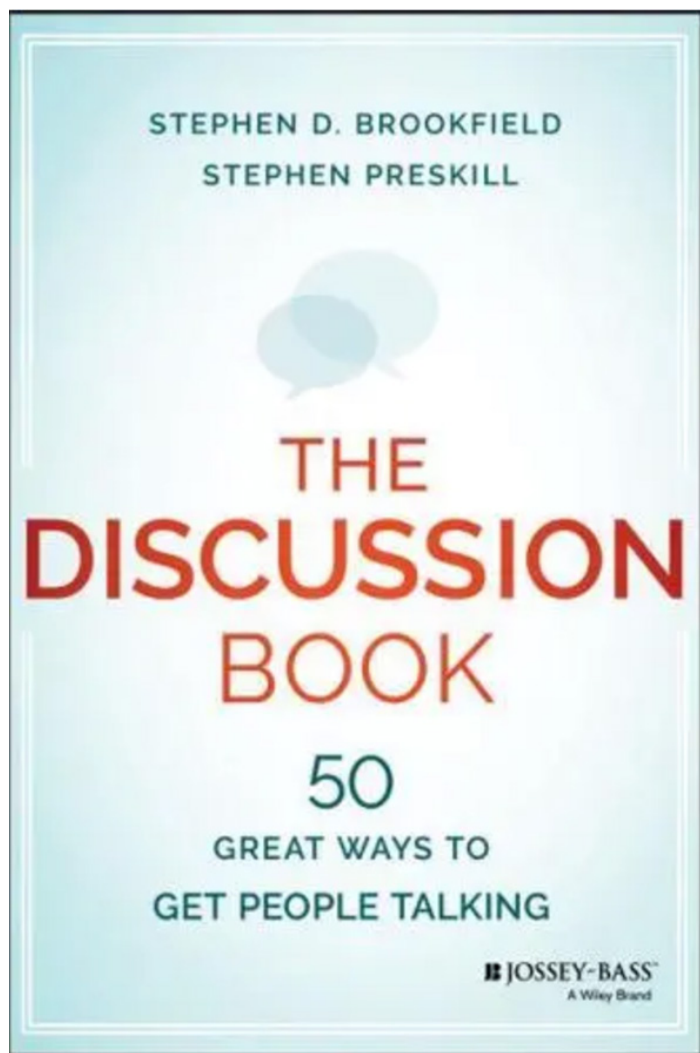


Figure 11. Book cover of *The discussion book* (Brookfield & Preskill, 2016).

I find that it's important as a teacher to set these protocols and explain the reasons for them, which usually are to slow things down, to give everybody a chance to think, to make sure that everybody has the opportunity to contribute at some point or another, to stop the power that students have because of their identity outside of the class just reproducing itself automatically inside, and to focus on raising questions and issues rather than coming up with specific answers or responses to problems. You constantly monitor what's going on by using backchannel chat, Slido, the Critical Incident Questionnaire or whatever classroom research device you want. You use all those things to monitor how things are going, and then you calibrate based on what you find out.

That would be the overall sequence of stages that I would employ in thinking about this. But having said that, some will be dropped, and some will be added, depending on context. But at least, as I go into a new situation, those are the things that are on my mind as design elements in terms of how I can sequence students' exposure to increasingly difficult ideas that threaten their sense of themselves as non-racist or anti-racist, 'good white people'.

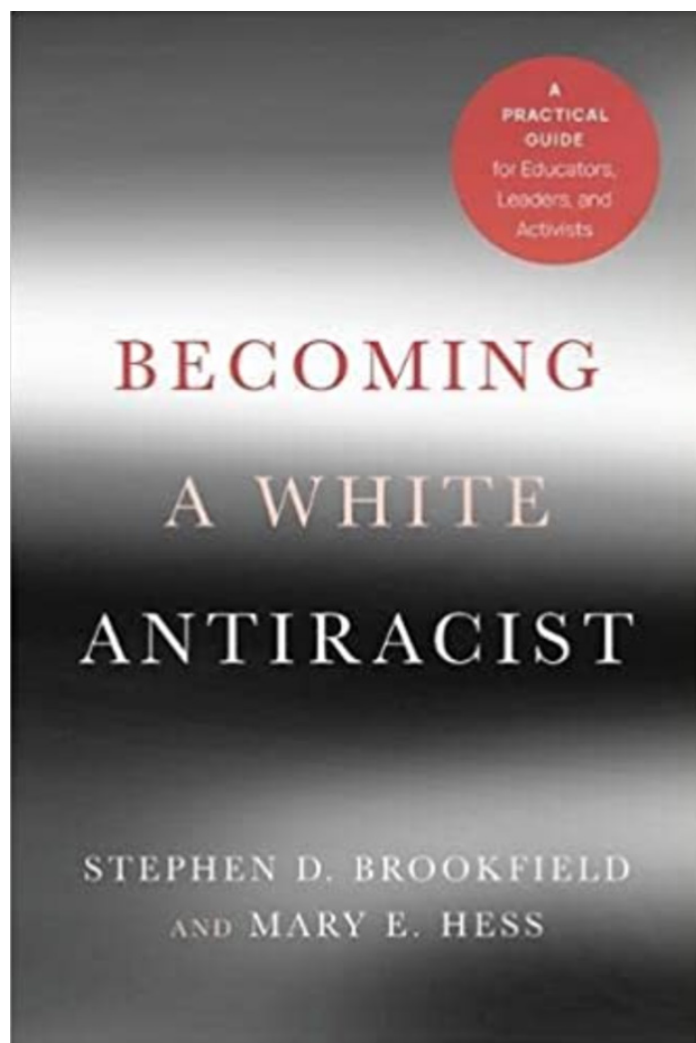


Figure 12: Book cover of Brookfield & Hess's *Becoming a white antiracist* (2021).

Second question: how can White teachers address the topic of race in ways that don't re-center their power? That is always a problem. Of course, as we know from my comments around power (see Brookfield et al., 2022, 2024), I don't pretend that teachers don't have power anymore because we do. But students do as well. For teachers, our question is: How do we use the power in a supportive, illuminating, and ethical way to help students' learning? One way that we can do that is by using a co-teaching model where the team comprises members who come from different racial backgrounds. Secondly, whenever you open yourself up to anonymous critique by all those classroom research approaches, Slido, CIQ [Critical Incident Questionnaire] etc., it opens your exercise of power to public questioning. When you report out comments that people have made, either questioning or supporting or challenging and condemning your use of power, when you talk about all that stuff as a public issue with a group, that helps create a dissent to your

power to some degree.

*I don't pretend that teachers don't have power anymore because we do typically. But students do as well. For teachers, our question is: How do we use the power in a supportive, illuminating, and ethical way to help students' learning?*

When white teachers use themselves as models of how they've learned racism and how they enact it, when we are the case study that we present to students, they see our own struggle and understand that we're constantly evolving and becoming and trying to get better at working in this area to understand the dynamics more fully and more accurately. I think that also using yourself as a case study of imperfection helps deconstruct your power as well.

Then the final sub-question: how can we not just understand racism but seek to change it? I do think that the first response I'm always considering is: 'What are the opportunities for us to act collectively? How do we build a network of people within the institution who are concerned about some institutional practices? If we're talking organizationally, how do we create alliances across different departments and different schools?'

At the university where I was employed, we had an anti-racist coalition of staff, students, lecturers and professors. It was about 250 people strong, drawn from every department, every unit or office in the university. Having that network meant that when the alliances wrote a letter to the president or when the alliance contacted the dean of a college about some policy or practice, it had some weight and authority behind it. Because you knew it was an organization that comprised 250 people and it would exercise collective leadership. The letter would speak on behalf of that big coalition.

Then, in the wider world, we know – at least it seems to me, that in order for a social movement to be more than a set of performative demonstrations, you really need some kind of political party or some organizational network: a permanent institution that is funded with permanent employees, and that focuses on the advancement of these issues. Setting up specific chapters of some national organization is the way, at least in the Civil Rights Movement, that change around race came. I still think that holds true very well. For example, the Black Lives Matter Movement has local affiliates, and the way it works looks different in Minneapolis than it does in New York City or Pensacola, Florida, or wherever.

The second thing on a more local level that I've emphasized a lot is if you don't have very much power and you are a junior member of the organization, always tie any project that you're pushing – any reform you're asking for, any new practice that you want to institute – tie all those things to the declared public mission of the institution or its value statement or its strategic plan or what it says it stands for. When you use the dominant language that's institutionally approved, and you couch the changes that you're suggesting or the issues that you're raising in that language, it's harder

for institutions to wriggle out of taking it seriously. It also protects your own status because you are just asking: 'How can we do better live out the values that we say we're all about?' So, that's my kind of long, omnibus answer to your question, breaking them down into those three sub-questions.

ST: This sequence that you were describing, though, of course, context-dependent, strikes me as extremely clever. The term 'brave space' is very interesting, and it is, of course, different from a 'safe space'. This could be quite controversial in Jürgen's and my more Confucian environment in Singapore. But you're immediately mentioning an alternative approach where you alert the students to the thought process. I like the part about the racial affinity groups because that seems particularly non-threatening.

SB: It's interesting that with the affinity groups where my colleagues and I use this approach, we typically get a lot of pushback and resistance from the white members who say, 'How are we going to learn about race if we're not talking to people of color about their racialized experiences?' I think that's a legitimate question to raise. But it then allows us to say: 'Well, in a white group, we can focus specifically on what it means to have a white racial identity. We can talk about our own sense of ourselves as having a racial identity or our own growing understanding of these issues. We can talk about all of these things in a way that it would be harder for us to do were we in a group with folks of color'. So, it's not as if affinity groups are usually welcomed. There's a lot of confusion about why we are doing this and how we will ever move forward if we don't talk about our differences and have a conversation around that. So, we say: 'Well, we're not saying 'don't do that', we're just saying that for some of the time, for some specific purposes, it's helpful to be in a racial affinity group'.

But that's something I should probably stress as well: when you do this as a white instructor with white students who are in a multiracial class environment, the white students will often resist it, and they'll have a hard time understanding why this is happening. But usually, when you then debrief the small group experiences in racial affinity groups, the members of color in their groups will talk about how refreshing it was to just meet with those of their own racial identity, to be able to relax, and not worry about how Whites are going to react to comments. I remember one group saying in a training we set up: 'We didn't have to massage white egos'. When people of color brought up an issue of race in this particular institution, Whites would become so alarmed because their notion of a 'good white person' was being challenged, and they would spend a lot of time defending themselves and trying to explain their conduct or their thinking. The people of color were always having to calculate: 'Well, how do we introduce this without making people feel threatened? How do we do it in a nice way?' 'My own future in this institution maybe is on the line if I'm too confrontational'. All those calculations for a person of color are much less important when you're in a group of others who are drawn from a similar racial background. So, hopefully, the white students hear those kinds of comments and feedback and get a better sense of why we're doing it. But it's often resisted earlier on.

Dealing with our own biases and stereotypical actions in the classroom

**JR:** In an article in 2014, you wrote, "In classes, I catch myself not challenging students of color and realize my so-called empathy, desire to be an ally, masks an embedded racist consciousness, which says, 'They can't take a strong challenge from a white person'" (Brookfield, 2014, p. 91). Could you discuss the circumstances where you realised your own biases and stereotypical actions? Are these related to being a 'good white person'?

**SB:** I have so many examples of circumstances where I've realized, and I'm continuing to realize, my own biases and stereotypical actions. One of the most important events was way back in the early 1980s when I started teaching at Columbia University's Teachers College in New York. There was one particular African American woman who came up to me after a class one day and said: 'Do you realize that when I speak, you never say anything, you just nod? I don't know what that means. Do you not say something because you don't understand what I'm saying? Is it too difficult to comprehend? Or is what I'm saying irrelevant, so you can't really connect it to what we're talking about?' She pointed out to me my tendency to stay silent. As your question suggests, that was an example of where I felt that Whites have had the stage for too long. I shouldn't allow my voice to be too dominant. In particular, if a student of color said something that I disagreed with, I felt like I couldn't really express that disagreement because that seemed like we were reestablishing power relations based on race. I assumed I would be acting in an authoritarian way and demonstrating that I'm not taking their experiences seriously. So, I concluded that I'd better not challenge even though I feel there's something inaccurate or misconceived about this particular contribution that someone's made.

I did a lot of that hanging back. It was because of students like the woman that I quoted – this was 40 years ago – that that really stuck with me. Then I remember, also in the '80s, someone came to me and said, 'There's some racism going on in the class between students; you need to deal with it'. And I said, 'No, that's not really my concern. These are all adult students; they can sort it out amongst themselves'. I deliberately evaded the emotional effort of dealing with a situation that I felt like I was not really qualified to deal with. I also had a history of deflecting the need to take race seriously before I started to understand the dynamics of power based on racial identities that played themselves out in classrooms and meetings. Eventually, I gave permission for my students or my colleagues to talk about those dynamics by, first of all, modeling and talking about them myself. People would say things to me that were quite disturbing. They would point out my microaggressions.

There was one where I was running a classroom discussion. I asked everybody to participate and give their opinions on the issue we were discussing. Everybody did, and then I started to sum up and pointed out similarities and differences in the comments. One of the white female students raised her hand and said: 'We haven't heard from another person in the group' who was a younger Asian-American woman. I couldn't believe that I had overlooked her. So, I apologized

and asked her to speak, and then over the break in the class, I was thinking to myself: 'How did I overlook her? I was sure everyone had spoken'. I realized it was a good example of microaggression, where you do something by not calling on a student. You make them feel invisible and ignored and not of value to you.

I went back to the group after the break and said: 'I think you just saw a really good example of a microaggression because I didn't mean to exclude this student. It just happened'. When something 'just happens' and seems natural and normal, and you're not even aware of what's gone on, those are the times in which dominant ideology is very prevalent. In this case, I really had no idea that I'd excluded someone, an Asian American woman. Then, some white students spoke up and said: 'Oh, you just had a moment of forgetfulness, don't punish yourself'. But then the Asian student spoke up and said: 'This has happened pretty much in every course I've taken at the university. I have felt constantly overlooked like no one is really interested in my opinion, and I don't think people really notice that I'm in the room'. I use that example a lot of something that happens all the time.

I know that I meet male gazes more easily than female gazes, I make eye contact with men more than with women, I tend to know the male names more than I know the female names, I tend to know white students' names or be more comfortable speaking them than some students of color where their names are just phonetically unfamiliar or difficult for me to pronounce. So, in order to avoid the embarrassment of pronouncing them wrongly, I don't call on the student because then I'd have to use that name. In all these little ways, in these micro-decisions you make in the middle of a class or meeting, you see your learned racism and this notion of 'well, the only important students really are the white students'. I'm horrified to think that I might believe that, but my actions sort of support that learned perception that comes from white supremacy.

That whole notion of a good white person has been incredibly influential on me, and I do credit Shannon Sullivan's *Good white people*, which was published in 2014. That was really influential on me, I read it and recognized a lot of what she was talking about in myself. I realized that this desire to be a good white person is part and parcel of white identity development. It's one of the things that happened earlier, probably in your development as a racialized person. If you're white, you cling to this idea that you're not one of the bad white people enacting racism. You're one of the good ones who somehow escaped it. You treat everybody the same irrespective of their racial background or the pigmentation of their skin. That awareness of being a 'good white person' has helped me understand the limits of a colorblind perspective; it's helped me understand the dangers of white 'saviorism' and a colonial approach where it's your responsibility to fix the problems of other people. This constant desire to prove how anti-racist you are to colleagues and students of color is all part of the good white person's identity.

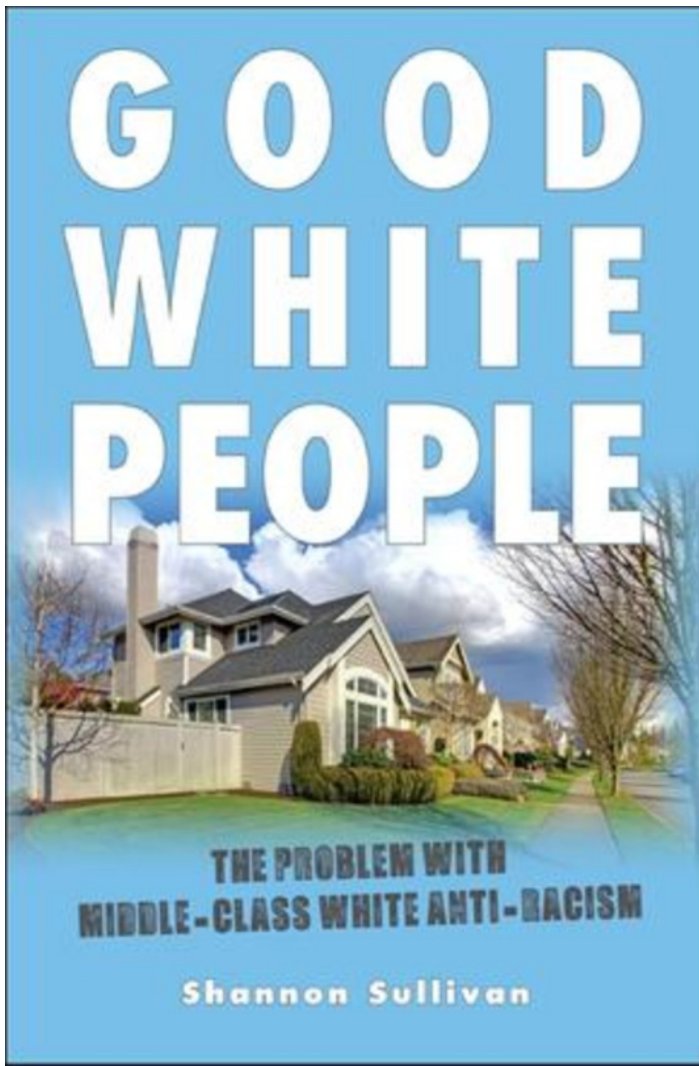


Figure 13: Book cover of Sullivan (2014).

*That awareness of being a 'good white person' has helped me understand the limits of a colorblind perspective; it's helped me understand the dangers of white 'saviorism' and a colonial approach where it's your responsibility to fix the problems of other people.*

If I didn't have several anonymous backchannels where students can point out things that they've noticed in Slido, backchannel chat and in the Critical Incident Questionnaire, it would be a lot harder for me to understand what was going on. So, that anonymous student's commentary and critique has also been really helpful over the years and students have said things about my own actions and my own words or decisions and choices in class. I've had to acknowledge that there's a great deal of truth in things that they've pointed out to me regarding my own behavior and view and that is an opportunity to model the kind of appropriate disclosure that you really need to do a lot of.

## How do we manage microaggressions and racism in a classroom setting?

ST: Would you like to further comment on the concept of microaggressions? In *Teaching race* (Brookfield & Associates, 2018), *The skillful teacher* (Brookfield, 2015) and other works, you highlight the usefulness of the concept of microaggressions. You wrote: "One of the most useful concepts I've stumbled across in the last few years has been that of racial micro-aggressions... [-] small acts of exclusion and marginalization committed by a dominant group toward a minority" (Brookfield, 2015, p. 119). Microaggressions are at the level of everyday behaviour that enacts the ideology of white supremacy and keeps racist systems in place. Microaggressions are defined as "daily verbal, behavioural, and environmental indignities, whether intentional or unintentional, that communicate hostile, derogatory, or negative racial, gender, sexual orientation, and religious slights and insults to the target person or group" (Sue, 2010, p. 5). Could you please share some examples from inside and outside of the classroom when you have witnessed, experienced, or enacted racial microaggressions?

SB: I've already dealt with that to some degree with the last question. I will add one, which is a dynamic that plays itself out over and over again in a mixed-race group: where a person of color points out some questionable comment that a white person has made, or a person of color or I point out how a particular action that someone has taken does have some embedded racism or white supremacy contained within it. When that happens, the other white members of the group will band together to save the white person whose behavior or comments are being questioned. I've seen this over and over again, where they'll say: 'Don't be so hard on him; he had a moment of forgetfulness, or he had a rough week'. Or: 'Not everything is about race; his actions have nothing to do with race'.

I've seen this happen in student groups, faculty groups, and in meetings of administrators where the other Whites have this informal pact to save a white person and explain away their behavior when a person of color identifies anything problematic. I try to name when that is happening, and I try to point out that it constitutes a micro-aggression. Your first response should not be to explain it away, justify your actions, and re-explain your real intent.

The first response is to acknowledge the harm that the other person feels has been caused. You don't try to deny or mitigate that harm by saying, 'Well, it wasn't really meant; you're being too sensitive'; you just let that person know that you hear what they're saying. Then, you apologize for the harm and take responsibility for it. Then maybe you can get to talking about what your intention was. But you don't start off by explaining your intention; you start off by acknowledging the harm and apologizing for your role in it. It's a pretty simple thing to say, and it is a very hard thing to do. Because Whites have this horror of being regarded as racist in any way. It all comes from this model of racism as something that's rooted in our individual psyche. That's why early on when you're talking about racism with a new group, you have to make sure, as best you can, that you're moving away from this individual understanding of it in terms of

your own individual moral failings and that you understand it as something that's systemically enacted, something that everybody learns and internalizes to a greater or lesser degree.

When I've been co-teaching with colleagues of color, when we finished our presentation or we've just done an activity, we ask people to debrief. Who did the questions go to? Do they go to me as the white male, so-called senior teaching member of the group, or do they go to the person of color or the woman? In most white spaces, it's typically clear that the questions tend to come to me because I'm assumed to have more authority and credibility in the teaching team. A lot of that is linked to my racial identity, and it's also linked to my gender identity. When that happens, we can point out as a teaching team that dynamic about how authority subconsciously is viewed as white and male, and that if someone has a white male body, they are assumed to contain more intelligence, have more credibility and be more competent.

Those are some of the typical dynamics that I've seen. Some of the things that I've done myself is to ask someone, 'Where are you from?' You see a student in class, and you assume that because they're not white, they're fairly recent immigrants. So, you ask, 'Where are you from?' and they'll say, 'Chicago' or 'Boise, Idaho'. Asking 'Where are you from?' implies that you're not from the United States. It just emphasizes the otherness, 'This is not a real American here'. 'This is an immigrant'. I should be particularly aware of that, given that I'm an immigrant myself, but I think my whiteness blinds me to the underlying message of that question that I have asked, 'Where are you from?' In my head, I'm just expressing simple curiosity, but of course, to the receiver, it's seen as, 'All right, you're not American, are you?' What other country or culture outside of this one do you come from?' Those are some examples of microaggressions and how I've tried to deal with them: acknowledge, apologize, take responsibility, and explain intent. I think the final typical stage when you're becoming aware that you've committed a microaggression is that you talk about how you're going to take account of what you've just learned and try to avoid replicating that behavior in the future. That's a final thing that we often teach you to do.

**JR:** How do you manage quite overt racism and white supremacy in a classroom setting? It's not just a microaggression but something a lot more blatant.

**SB:** I think you have to point it out. One of the things that students of color have made very clear to me as a white teacher: they've told me that if I let stuff go unchallenged such as a racist comment that somebody has said or posted on Slido, I have to acknowledge it. Students tell me that if I don't acknowledge what is going on and tell them how I'm going to try to address it, then they really don't have any respect for me. It's very hard for me to trust my own commitment to antiracism. Even though I'm a non-confrontational person by cultural training, I know from experience that when an overt expression of racism or white supremacy happens, I have to acknowledge it right then and there at the moment and bring it to the attention of everyone and talk about it. However, the way I respond to that will partly depend on

my reading of the context because if I'm saying we all have racism within us, it will be no surprise when it comes out. In fact, it would counter my own understanding of racism to always jump on that and say, 'bad white person'.



We'll start at 9.00.am. (EST)

Please log into **sli.do** & enter code

**4864957**

We will use this anonymous backchannel throughout the 2 days to ask questions, raise issues, make critiques & give reactions

Figure 14: Screenshot of an announcement regarding the use of Slido during an intensive weekend seminar. *Learning as a way of leading* (see Preskill & Brookfield, 2009) was co-taught by Stephen Brookfield and Stephen Preskill in May 2022 at Columbia University's Teachers College.

One of the things I'll often do is when I hear or see something, I'll say, 'I think what you just talked about is a really helpful example of learned white supremacy or learned racism'. Sometimes, I'll say, 'How would you feel if that comment was directed at you yourself, based on your racial identity?' Sometimes I'll say, 'Let's just stop for a moment. How do you think that comment was heard by people of color in this class, or how do you think that comment would be heard outside of this class by a stranger or by a person of color that you knew?'

Another approach I'll often enact is to say: 'What you've just said sounds so much like me. I remember thinking that and saying that and doing that. Then, what gave me pause and made me understand that there may be some racism embedded within it was such and such'. Using myself as a case study or an example sometimes takes the pressure off the person whose behavior has been identified and pointed out. When I'm doing this, I always try to give an example of a different formulation of words or how a question might have been posed that would have felt like a less racist or micro-aggressive kind of comment or question.

But if something really overt happens, like an attempt to shut someone down or to belittle them because of their racial identity, then you just have to name that and say that's not acceptable. 'That is not what we're about in this class and I have fairly precise participation grading rubrics that I use. A comment like that would clearly be a contravention or contradiction of a lot of the ground rules that I or the group have developed as examples of good participation. So, you can point that out. Sometimes I've just said 'I really feel that's an example of the worst kind of racism that we have to deal with. And while you might not have meant it that way, it constitutes racial stereotyping and dismissal'. Racial stereotyping is very common, and you just have to sometimes be very explicit in pointing out what's going on.

### **Developing an antiracist white identity**

**ST:** You wrote: "I came to understand that education about racism was often done best through narrative disclosure" – and you've spoken about that, of course, already – "not just through sharing tips and techniques of what does, or doesn't, work in confronting racism" (Brookfield, 2015, p. 113). Could you elaborate on how self-disclosure helps people develop an antiracist white identity?

**SB:** I have seen a certain dynamic play itself out in quite a bit of antiracist training over the years. That is where often a white person will come in and present themselves as a fully-formed white antiracist, whose job it is to enlighten other people who are earlier in that journey and then to go straight into the teaching or the training around antiracism. I always feel like that's a fundamental mistake: that if you come in, the first thing you as a facilitator have to do is to model your own experiences of racism, the times when you've enacted racism. So, give lots of examples from your own life and emphasize that this is something that you're still struggling with. Coming in as a living example of someone who's still trying to learn about this and who is not at the end of their journey by any means sets a tone for a workshop for other Whites who might be afraid of admitting to anything that is not politically correct or an example of a fully formed antiracist person.

If you just come in and talk about how you, as the facilitator or teacher, are just as ensnared in the system as anybody else – that I think is a helpful way of opening people up and setting a tone. I'm always trying to give examples of actions I've taken, decisions I've made or things that I've said. I have learned a certain way of understanding the world, that was internalized very early on, supported throughout my teenage and adult years, meaning that racism and white supremacy are just baked into the cake of daily life (I'm not sure that's a very good metaphor). They're basically learned; you're not born thinking these things. But if you're in a white supremacist culture, it would be crazy not to have some of that within you or to feel that you had somehow escaped it by a fierce moral commitment not to be racist. I'm always trying to use the example of my own narratives to teach people that racism is something that's structural, that it's internalized, and that everyday institutional practices and policies support – without us knowing it – the ideas of white supremacy.

Let's move away from this racism-as-an-individual-moral-failing model and see instead racism as something that's culturally learned. If you can get that understanding across at some basic level, people find it easier to deal with this. Because you're now not blaming individuals for their moral failings, you're just saying: 'Of course, it would be very strange if you didn't think or act this way, given the culture that you've grown up in'. Also, you have more success in getting people to develop an antiracist white identity and even to understand what being a white person means if you start off with narratives rather than statistics or theories. I've seen a lot of workshops start off with tables, representing massive inequities of access to health care or education, or disproportionately high numbers of inmates of color in penal institutions. Those are obviously important.

However, what engages people initially is a story that they recognize, and they can place themselves in that story and think, 'Yeah, I've done something like that or close to that, or I can see how I would do that in that situation'. When you use your examples from your own narrative experience, it connects with people in a way that statistics or fierce polemic about the need to be antiracist doesn't. I feel that you should bring in the stats and the studies after there has been some initial narrative disclosure. You're going to have more success in developing an antiracist identity in that way if you start talking in personal terms rather than in general or abstract terms. Finally, given that this work is strongly emotional, having a model of someone who's in the role of facilitator or leader talk about their own emotional responses and their own confusion, fatigue and frustration, is a very helpful way of bringing people around to thinking through: What does it mean to have a white identity? What does it mean to have an antiracist identity?

### **'Failing well' in antiracist workshops**

**JR:** We were surprised to learn that you conduct antiracist workshops for participants whose attendance is compulsory, so they may not be there out of their own free will. Could you share the reasons why you expose yourself to such (presumably at least partially painful) experiences and share some of them with us?

**SB:** It's a question that most people who are involved in staff training, faculty development, or professional development ask themselves: Should this be mandatory or voluntary? Of course, when that kind of training is voluntary, you tend to preach to the choir of selected individuals who see the importance and the necessity of the work. I'm somebody who's done a lot of that kind of development myself. One part of me says: 'It was great that these people are here, and it's important that they are fueled in that desire to be good teachers and that they see there are others in the institution who share their passion and desire for improvement. They need support.' But I'm always thinking, 'These are not really the people that should be here'. The people who should be here are those who dismiss the notion that there is anything to improve in their practice and who feel that any kind of challenge to their ideas or habitual ways of acting is disrespectful in some way to them. If you think about how change happens institutionally or organizationally, it's



always a collective change.

Selected individuals, a new president or a new principal of a school can set a tone. But ultimately, if an institution is going to change, it has to have some collective engagement in intentional change across the whole institution. So, that's why I feel that sometimes we need to have mandatory training. It is a very different dynamic because there'll be a lot of skepticism and anger in the room. There'll be attempts to sabotage what's going on by calling your credibility into question, and there'll be constant dismissing of your authority. But when you're confronting people with a picture of the world that's 180 degrees different from the one that they thought they were walking in every day, then you have to expect that you get all those forms of pushback and resistance when you mandate it. But it's important to remember that when people respond that way, criticize you and call your authority into question, you shouldn't take it personally. It feels like it's personal, but whoever was running that mandatory education or training effort would have the same criticisms levelled at them. It doesn't matter who's doing it on one level and who's in charge of it because just the instigator of that antiracist education is going to be seen as the enemy by a lot of people. It just comes with the territory. You have to try to depersonalize all this criticism and realize it's not directed at you personally.

This is one of the benefits of doing it as part of a team. When you have team-facilitating mandatory training, you have built-in support there from your team members who can tell you you're not crazy and who can point out good things that happened when you felt you failed miserably. They can give you a different reading of how a meeting went or point out things you hadn't noticed in a classroom. Then, if you really have got into a difficult situation, usually the team members can talk about how they have been in exactly the same situation, help talk you down, and talk you through it. Doing mandatory training on your own consistently can be pretty wearing and demoralizing. So, it's another reason why I am always going for the team approach.

**ST:** You provided an excellent quote earlier: that you can do this kind of antiracism training or education either imperfectly or not at all. Our next question is along the same lines: What does being successful in teaching race mean? Is the best possible outcome to 'fail well' (to cite Samuel Beckett)? Could you elaborate on some of the most important misperceptions that block white teachers' efforts to do antiracist work? In *Teaching race*, you discussed the following eight avoidable mistakes: 'I can control what happens', 'I need to stay calm', 'I must fix racism and transform my students', 'I've finally escaped racism', 'I understand your pain', 'Please confess your racism', 'I mustn't dominate, so I'll stay silent', and 'I'm your ally'.

**SB:** Again, I'll break these down, taking each of the sub-questions that you raise. You start off with: 'What does being successful in teaching race mean?' I've already talked a lot about how we have to readjust our notion of what success looks like in this work. Unless we do that, we're going to feel constantly as if we're incompetent and we have no idea what we're doing. You'll start to doubt yourself. So, you have to understand that success is not all the things we talked about:



Figure 15. Photo of Samuel Beckett in 1977 by Roger Pic, Bibliothèque Nationale de France, public domain. Failure is a central theme in Samuel Beckett's oeuvre. A famous passage from *Worstward ho* (Beckett, 1989, p. 101) reads: "Ever tried. Ever failed. No matter. Try Again. Fail again. Fail better."

staying calm, coming to consensus, leaving a classroom, feeling everybody has had their perspective transformed, or avoiding upsetting expressions of emotion. We have to leave all that behind. For me, I start thinking about whether people are willing to come back and continue the conversation. That is the key criterion of success that I use: whether or not someone is willing to keep talking about this or keep trying to push back against institutional practices and try out new policies and new ways of admitting students, assessing students' learning, or appointing and promoting people with institutions. If we're still willing to continue talking and acting in that way, then that is just success in itself.

A lot of this what we're facing is really a Eurocentric viewpoint, privileges perfection and seeks constantly for the correct way to do things. It's a constant binary emphasis you sometimes see in Western thought: 'There are best practices, and there are worst practices'. 'There are effective criteria or effective approaches, and then there are ineffective approaches'. That Eurocentric epistemology really does get in the way. You just got to start thinking about, 'Well, I'm going to do it imperfectly or not at all, and those are your only two options'. Having this Eurocentric notion that I can become a really good, certified trainer of this who is mistake-free or a certified teacher in this area is just the wrong way of thinking about it. Because, quoting Beckett, you will 'fail' if that's how you are assessing the effectiveness of your work. In my own experience, I always want to go back and do a particular session or training over again. I've realized, after a lot of years, that that's just the nature of the work. I will

always feel like 'I could have done that better. Man, I wish I'd said this or done that at that particular point'. And I'll say to myself, 'I really missed an opportunity there'. But I just have to understand that this is so complex that it's always going to be part and parcel of the work. The most you can do is to understand better why the particular dynamics that you were dealing with were in play.

In terms of some of the most important misperceptions, we have talked about some. But I'll just say something briefly on some of these. The 'I can control what happens'. Well, you can't; you can plan, you can learn from experience, and you can go in with the sort of sequence that I outlined earlier. But one thing you can depend on is that something is going to take you by surprise; some new dynamic or some new manifestation of an old dynamic is going to emerge. You just need to be prepared, knowing that that is almost certainly going to happen, and not feel uncomfortable by calibrating and changing plans in midstream, by being flexible, and by adjusting to what you're learning about a group. I think if you just talk out loud about that process of decision-making and how you're interpreting what's going on in the class, that constitutes a good model of a critically reflective and responsive practitioner.

This emphasis on keeping calm: 'I need to keep things calm, I need to stay calm' is what Bell Hooks (2014) called "bourgeois decorum" as the model of conversation in higher education classrooms. But racial discussions won't stay calm, there will be raised voices and tears and expressions of anger and heated conversation, lots of awkward silences. But the silence is often just a necessary pause for people to process and mull over some very complex information or ideas that they just encountered. So, all that stuff is normal, it's not a sign that things have gone off the rails or gone wrong. That's something again that a lot of my colleagues, plus myself, have had a hard time understanding because we would like everything to work out the way that we've anticipated. When it doesn't, it's so easy to think, 'Oh, I've lost the plot. I've got it wrong. I haven't planned properly, and I'm an imposter. I don't deserve to be doing the work'.

It's easy to slip as a white teacher into this role of thinking, 'Well, I am the racially cognizant one. I'm going to bring the rest of you unenlightened, unsophisticated people into a more enlightened state of being'. If you have that attitude, people pick that up very quickly. It really puts them off. Plus, it takes a long time to have a significant personal change. It's not something where you go to a workshop, and you say, 'Oh wow, there's this whole other way of living'. Then from that point onward, you're engaged in this other way of living, according to this completely different paradigm of how the world works. It doesn't happen like that. It's a lot of halting moves forward and then regression to earlier behaviors and then moving forward again when you feel you have a bit more courage, then regressing again when things are difficult.

Saying something like, 'I finally escaped racism' is all to do with presenting yourself as a fully formed antiracist. But no one has escaped racism. People have done better and may be aware of how racist practices and instincts are shaping the decisions and actions in their lives. But they haven't



Figure 16. Gloria Jean Watkins, better known as Bell Hooks (1952 – 2021), at the New School, 10 October 2014; photo by Alex Lozupone; CC BY-SA 4.0. Renowned for her insights on race, feminism, and class, Bell Hooks served as a Distinguished Professor at Berea College. She authored around 40 influential books, including essays, poetry, and children's books, delving into the intersection of race, capitalism, and gender, and their role in sustaining oppression and class domination.

escaped it because the system is so deep and endemic. They might be able to negotiate their way within it. But they will never have escaped it. So, even as you're challenging the ideas of white supremacy, it's still going to be an important framework that's determining your actions.

### *No one has escaped racism.*

'I understand your pain' - I've seen this spoken by white students in multiracial groups as a response to an expression of being on the receiving end of racism that comes from a member of color in the group. The whites will often try and say: 'I know what you're experiencing and where you're coming from. I was in this situation once in my life, people laughed at me because of my working-class accent, and I was considered unintelligent because of the way that I spoke. Usually, attempts by Whites to say, 'I understand your pain' by drawing a connection between experiences of racism and their own experiences do not go well. It sounds to people of color that you're trying to diminish their own experiences.

I do feel that when you're looking at processes of marginalization, it is appropriate to make a link between the way that you as a white person have felt marginalized, or the way that you as a woman or a trans-person have been marginalized. I do understand that is one way of talking about this dynamic within a mostly white group. But it's a major mistake for Whites to say, 'I understand what it's like to be on the receiving end of sustained racism every day of your life' because we can't really understand that.

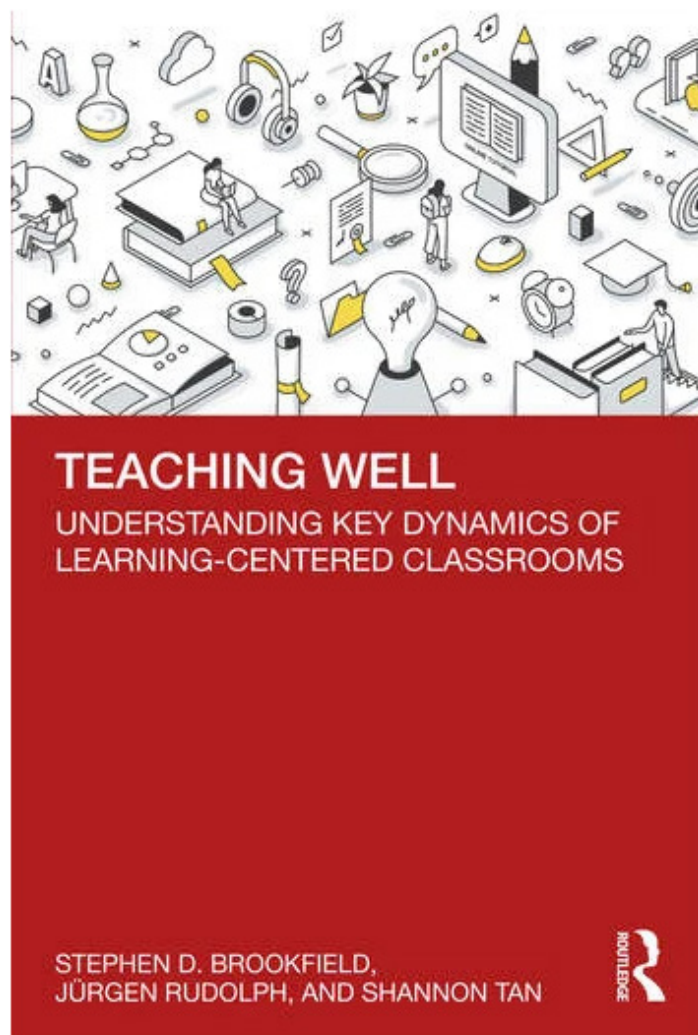


Figure 17. Our book on teaching well (Brookfield et al., 2024). This interview is a significantly expanded version of Chapter 9. For reviews of the book, see Xhemaili (2023), O'Brien (2024), Waring (2024) and Day (2024).

The 'confessing your racism' dynamic is where you're running a class or doing a training that's supposed to make people antiracist and participants will spend a lot of time confessing to their past sins of racism. They'll use those confessions as a sign of how woke they now are. While, again, I've argued constantly that it's important that people express their own narratives and talk about their own stories and their own experiences, you have to make, as an educator, leader, or trainer, sure that confessing to those experiences doesn't dominate the whole workshop – where white students are in effect turning to students of color and saying: 'Please absolve me from my sin [all laugh]. Please tell me I'm a good white person. And tell me 'No, that was you in the past, but you're not like that anymore', 'We consider you an ally.' and so on.

One of the things that I hear a lot from white students or white colleagues is essentially: 'Well, I can't really contribute to this. I certainly can't exercise any leadership in this area because I have no experience of race. I don't know what it's like to be on the receiving end of sustained racism every day of your life'. That's clearly accurate. But I always say to a white group: 'Well, yes, you don't have that experience of racism, but you do have an enormous experience of race from the perspective of an unconscious enactor of it, someone who knows how these ideas are transmitted and learned'. Also, as someone who knows how easy it is to move through institutions and communities and families every day and not to be aware of the racism that is favoring and advantaging you in some way. If you assume that you can't do any leadership on this because you don't have a racial identity, that race is something only people of color have, that you haven't been the recipient of all these racist policies – well, I will say: 'You actually have been the recipient of racist policies, but what you've experienced on the receiving end is a removal of barriers. Because you don't have these barriers and these stereotypes weighing down on you, you'll never have to think about them. So, please don't feel that you don't have an experience of race. You do. It's just from a very different perspective. The more you come to understand that your white identity is the norm for what correct thinking or correct behavior or normal, universal behavioral thinking looks like, and the more you understand how that's happening, the more you're able to teach about the dynamics of white supremacy'.

The final thing is not to go around saying 'I'm your ally', or 'I want to be your ally'. Act as if you are, try and do that. But don't announce it because that's a very performative act. Really, you're looking for absolution, and you're looking for evidence that you're a good white person and that you're not a racist. I'll always say, 'Well, yeah, it's good to understand what it means to act as an ally or an accomplice or a co-conspirator, particularly in movements or in specific projects run by folks of color'. So, be an ally, but don't declare it as a self-identifier. Because then you will not be taken seriously in my experience by the folks of color that you're working with. However, if someone calls you that, then you should feel justifiably honored – but don't come in saying that you are; it's a fundamental mistake.

JR: Thank you so much for the interview, Stephen!

### Acknowledgement

We thank Dr Begüm Burak for her excellent transcription of the interview.

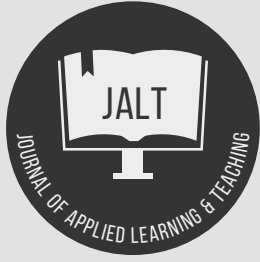
### References

- Barnett, P. E. (2018). Building trust and negotiating conflict when teaching race. In S. D. Brookfield & Associates (2018). *Teaching race: How to help students unmask and challenge racism* (pp. 109-130). Jossey Bass.
- Beckett, S. (1989). *Nohow on*. John Calder.

- Bell, D. A. (1995). Who's afraid of critical race theory? *University of Illinois Law Review*, 893.
- Brookfield, S. D. (2005). *The power of critical theory: Liberating adult learning and teaching*. Jossey-Bass.
- Brookfield, S. D. (2014). Teaching our own racism: Incorporating personal narratives of whiteness into anti-racist practice. *Adult Learning*, 25(3), 89-95.
- Brookfield, S. D. (2015). *The skillful teacher: On technique, trust, and responsiveness in the classroom*. John Wiley & Sons.
- Brookfield, S. D. (2017). *Becoming a critically reflective teacher* (2nd ed). John Wiley and Sons.
- Brookfield, S. D. (2018). The dynamics of teaching race. In S. D. Brookfield & Associates (2018). *Teaching race: How to help students unmask and challenge racism* (pp. 1-17). Jossey Bass.
- Brookfield, S. D. (2023). Foreword. In A. Darder, C. Hayes II, & H. Ryan (Eds.), *On class, race, and educational reform: Contested perspectives* (pp. xvii-xxi). Bloomsbury Academic.
- Brookfield, S. D., & Associates. (2018). *Teaching race: How to help students unmask and challenge racism*. Jossey Bass.
- Brookfield, S. D., & Hess, M. E. (2021). *Becoming a white antiracist: A practical guide for educators, leaders, and activists*. Stylus.
- Brookfield, S. D., & Holst, J. D. (2010). *Radicalizing learning: Adult education for a just world*. John Wiley & Sons.
- Brookfield, S. D., & Preskill, S. (2012). *Discussion as a way of teaching: Tools and techniques for democratic classrooms*. John Wiley & Sons.
- Brookfield, S. D., & Preskill, S. (2016). *The discussion book: 50 great ways to get people talking*. John Wiley & Sons.
- Brookfield, S. D., Rudolph, J., & Tan, S. (2022). Powerful teaching, the paradox of empowerment and the powers of Foucault. An interview with Professor Stephen Brookfield. *Journal of Applied Learning and Teaching*, 5(1), 131-145. <https://doi.org/10.37074/jalt.2022.5.1.12>
- Brookfield, S. D., Rudolph, J., & Tan, S. (2024). *Teaching well: Understanding key dynamics of learning-centered classrooms*. Routledge.
- Brookfield, S. D., Rudolph, J., & Yeo, E. (2019). The power of critical thinking in learning and teaching. An interview with Professor Stephen D. Brookfield. *Journal of Applied Learning and Teaching*, 2(2), 76-90. <https://doi.org/10.37074/jalt.2019.2.2.11>
- Colin III, S. A. J., Sheared, V., Johnson-Bailey, J., & Brookfield, S. D. (2010). Epilogue. Implications for curriculum, programming, and research. In V. Sheared, J. Johnson-Bailey, E. Peterson, S. A. Colin III, S. D. Brookfield & Associates (2010). *The handbook of race and adult education: A resource for dialogue on racism* (pp. 343-374). Jossey Bass.
- Cunningham, P. M. (2010). Foreword. In V. Sheared, J. Johnson-Bailey, S. A. J. Colin, III, E. Peterson, S. D. Brookfield & Associates (Eds.), *The handbook of race and adult education* (pp. xxv-xxviii). Jossey-Bass.
- Day, R. (2024). Book review of Brookfield, S., Rudolph, J., & Tan, S., "Teaching well. Unveiling the art of learning-centred classrooms". *DBS Applied Research and Theory Journal, March*, 106-107, <https://journal.dbs.ie/index.php/dbs/article/view/125/62>.
- Draper, J. (2015, October 21). No Irish, no blacks, no dogs, no proof. *The Guardian*. <https://www.theguardian.com/money/2015/oct/21/no-irish-no-blacks-no-dogs-no-proof>
- Galtung, J. (1969). Violence, peace, and peace research. *Journal of Peace Research*, 6(3), 167-191.
- Hess, M. E. (2018). Using digital storytelling to unearth racism and galvanize action. In S. D. Brookfield & Associates (2018). *Teaching race: How to help students unmask and challenge racism* (pp. 253-272). Jossey Bass.
- Hooks, B. (2014). *Teaching to transgress*. Routledge.
- Klein, M. (2018). Teaching intersectionality through "I am from..." In S. D. Brookfield & Associates (2018). *Teaching race: How to help students unmask and challenge racism* (pp. 87-108). Jossey Bass.
- Marcuse, H. (1969). Repressive tolerance. In R. Wolff, R., B. Moore, & H. Marcuse (1969). *A critique of pure tolerance*. Beacon Press.
- O'Brien, J. (2024). Book review of 'Teaching well: Understanding key dynamics of learning-centered classrooms' by Stephen D. Brookfield, Jürgen Rudolph, and Shannon Tan. *Learning: Research and Practice*, 10(1). <https://doi.org/10.1080/23735082.2023.2300440>
- Oluo, I. (2018). *So you want to talk about race*. Seal Press.
- Petcher, A. (2012). *Robertson's jam and Golly badges*. Blog post. <https://aipetcher.wordpress.com/2012/08/23/robertsons-jam-and-golly-badges/>
- Preskill, S., & Brookfield, S. D. (2009). *Learning as a way of leading: Lessons from the struggle for social justice*. Jossey-Bass.
- Rudolph, J. (2019). Book review of Brookfield, S. D. (2017). *Becoming a critically reflective teacher*. Second edition. *Journal of Applied Learning & Teaching*, 2(2), 122-124. <https://doi.org/10.37074/jalt.2019.2.2.22>
- Rudolph, J. (2020). Book review of Brookfield, S. D. & Associates (2019). *Teaching race. How to help unmask and challenge racism*. *Journal of Applied Learning & Teaching*, 3(1), 151-156. <https://doi.org/10.37074/jalt.2020.3.1.21>

- Rudolph, J. (2022). Book review of Brookfield, S. D., & Hess, M. E. (2021). *Becoming a white antiracist*. *Journal of Applied Learning & Teaching*, 5(1), 203-207. <https://doi.org/10.37074/jalt.2022.5.1.23>
- Said, E. W. (2019 [1978]). *Orientalism*. Penguin Classics.
- Sheared, V., Johnson-Bailey, J., Colin III, S. A. J., & Brookfield, S. D. (2010). The beginning. Kitchen table dialogue. In V. Sheared, J. Johnson-Bailey, E. Peterson, S. A. Colin III, S. D. Brookfield & Associates (Eds.), *The handbook of race and adult education: A resource for dialogue on racism* (pp. 1-25). Jossey Bass.
- Sheared, V., Johnson-Bailey, J., Peterson, E., Colin III, S. A., Brookfield, S. D., & Associates (2010). *The handbook of race and adult education: A resource for dialogue on racism*. Jossey Bass.
- Smith, B. (2018). Forming classroom communities to help students embrace discomfort. In S. D. Brookfield & Associates (Eds.), *Teaching race: How to help students unmask and challenge racism* (pp. 171-189). Jossey Bass.
- Sue, D. W. (2010). *Microaggressions in everyday life: Race, gender, and sexual orientation*. John Wiley & Sons.
- The Black and White Minstrel Show* (n.d.). <http://www.turnipnet.com/whirligig/tv/adults/other/tbawm.htm>
- Upton, B. (2023). *The adventure of two Dutch dolls and a 'Golliwogg'*. Good Press.
- Waring, P. (2024). Book review. Brookfield, Stephen D., Rudolph, Jürgen & Tan, Shannon (2024). Teaching well: Understanding key dynamics of learning-centre classrooms. Routledge. *Journal of Applied Learning and Teaching*, 7(1), 1-2. <https://doi.org/10.37074/jalt.2024.7.1.5>
- Xhemaili, M. (2023). Book review of 'Teaching well. Understanding key dynamics of learning-centered classrooms'. *Thesis*, 12(2).

Copyright: © 2024. Stephen D. Brookfield, Jürgen Rudolph, and Shannon Tan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



---

## The use of Generative AI in qualitative analysis: Inductive thematic analysis with ChatGPT

---

Mike Perkins<sup>A</sup>

A

*Centre for Research & Innovation, British University Vietnam*Jasper Roe<sup>B</sup>

B

*School of Social and Health Sciences, James Cook University Singapore*DOI: <https://doi.org/10.37074/jalt.2024.7.1.22>

---

### Abstract

This article describes a methodological innovation in the analysis of qualitative data using Generative AI (GenAI) tools alongside traditional research methodologies to conduct inductive thematic analysis. The case study employs an integrative method that comprises two researchers conducting simultaneous analysis: one using manual and traditional research approaches to coding, analysis, and interpretation, and the other conducting the same analysis but with the support and assistance of GenAI tools, namely, the premium version of ChatGPT (GPT-4).

The key strengths of this approach include the enhanced capacity for data processing and theme identification offered by GenAI, along with the nuanced understanding and interpretative depth provided by human analysis. This synergy allows for a richer and more complex understanding of the themes present in the data. The challenges encountered include managing the inconsistencies and hallucinations of GenAI outputs and the necessity for rigorous validation processes to maintain research validity. The findings indicate a complementary relationship between GenAI and human researchers, where the use of such tools can expedite the analytical process without diminishing the essential role of the researcher's expertise and critical engagement.

**Keywords:** AI research methods; ChatGPT; Generative AI; inductive thematic analysis; qualitative analysis; research methodologies.

### Introduction

The development of Generative Artificial Intelligence (GenAI) has had a major impact on a range of sectors, with the potential for significant reshaping of numerous sectors, from medical diagnostics (Caruccio et al., 2024) to financial services (Dwivedi et al., 2023). Academia currently finds itself at a crucial turning point as AI capabilities undergo

exponential growth, doubling in intricacy every half-year (Pichai, 2023), with significant implications for how teaching and learning may occur in this new landscape with the existence of multiple paradoxes (Lim et al., 2023). In the context of educational research, the advent of foundation models (FMs) such as Gemini and GPT-4 signifies more than technological innovation; it is a transformative moment for academic research methodologies. Attracting a broad spectrum of stakeholders (Bowman, 2023), applications built on top of FMs such as ChatGPT (OpenAI, 2022), promise a potential for enhancing research capabilities. From aiding Non-Native English Speakers (NNES) in articulating complex academic ideas (Roe et al., 2023) to developing new methodologies in personalized learning research (Kasneji et al., 2023) and potentially improving opportunities for student-teacher collaboration and communication (Limna et al., 2023), their potential is vast. However, their very sophistication raises questions about academic authenticity and integrity (Abd-Elaal et al., 2022; Cotton et al., 2023; Köbis & Mossink, 2021; Perkins, 2023).

Furthermore, for academic researchers in education and social sciences, these GenAI tools offer a challenge to overcome. On one hand, they can expedite research processes and provide alternative perspectives through AI-generated content. However, they require a rigorous re-evaluation of methodological standards and research integrity, especially given the problems of discerning between AI-generated and human-authored content (Anderson et al., 2023; Chaka, 2023; Elali & Rachid, 2023; Hassoulas et al., 2023; Perkins, Roe, et al., 2023; Weber-Wulff et al., 2023). As these tools improve in their capability, and new AI tools emerge and become central to academic research, a similar evolution is occurring at the student level. Students are increasingly using these tools, either to support their learning, or for the unethical completion of required academic assessments as an instrument of potential misconduct, sparking academic integrity concerns in higher education (Sullivan et al., 2023). Consequently, tools have been developed to detect AI-produced writing, but research has shown that these

are neither effective nor ready for deployment (Chaka, 2023). This underscores the need for academia to not only understand and integrate these tools responsibly but also to train students in their ethical and effective use to ensure the responsible propagation of GenAI tools in tomorrow's research landscape. To this end, new methods of integrating AI into education are emerging, such as AI-enabled assessment (Perkins et al., 2023) and encouraging open educational practices (Mills et al., 2023). In this context, it is important to find a balance between taking advantage of the powers of GenAI, and equally ensuring it is used responsibly and ethically, while being aware of both the risks and benefits of GenAI in research (Salah et al., 2023).

### **Using Foundation Models and ChatGPT for research purposes**

As GenAI tools become more integral to the research process, there is a need to prepare the next generation of researchers to use these tools, which goes beyond tool familiarization. Further research is also needed to understand how scholars currently perceive the role of these tools, and empirical studies conducted so far have shown that opinions vary, ranging from 'changing the role of educators' to 'extending the human brain' (Firat, 2023). Institutions must ensure students grasp the ethical, methodological, and practical implications of GenAI integration into research (Foltynek et al., 2023). Such foundational knowledge will be instrumental as students transition to leadership roles in academic research. Such advancements empower researchers, offering tools that streamline everything from preliminary data explorations to intricate statistical model constructions (Baidoo-Anu & Owusu Ansah, 2023). Yet, with power comes responsibility; the nuances of these tools, their underlying biases, and potential pitfalls must be rigorously understood and navigated.

Currently, FMs act as the foundation for a variety of GenAI applications. ChatGPT by OpenAI was one of the first GenAI tools to gain widespread recognition, largely owing to its fluency in language capabilities and intuitive chatbot style instruction. Initially designed for tasks such as text generation and dialogue, the tool's capabilities were expanded to include additional analytical functions with the introduction of Advanced Data Analysis mode (OpenAI, 2023), formerly known as Code Interpreter. This enhanced functionality allows ChatGPT to perform additional tasks beyond simple text-based interactions, such as data analysis, complex problem solving, and code generation. These capabilities make ChatGPT an important tool for academic inquiry across various disciplines (Perkins & Roe, 2024; Salah et al., 2023) because of its capacity to swiftly review extensive datasets, discern patterns, and provide nuanced interpretations complemented by traditional human-centric analytical approaches (Bowman, 2023). Such synergy between machine and researcher can lead to richer insights and more comprehensive interpretations. However, challenges are present. The stochastic nature of these tools can pose hurdles to result replicability, which is a key foundation for research validity and integrity. Additionally, the requirement to create effective prompts requires dual expertise from researchers: command over the subject

matter and a nuanced understanding of AI-driven analyses (Roe et al., 2023). Significant ethical implications must also be addressed, as they relate to publication of ChatGPT supported publications (Lundgren et al., 2018; Rahimi & Abadi, 2023; Xames & Shefa, 2023).

### **Using Foundation Models for qualitative data analysis in educational research**

These tools have been demonstrated to streamline processes, from supporting in interviews (Chopra & Haaland, 2023), translating content (Chen, 2023) and manuscript preparation (Xames & Shefa, 2023; Zhai, 2022). Some researchers have highlighted the limitations that GenAI tools have with regard to supporting with the literature review process (Haman & Školník, 2023), but newer versions of these tools are significantly more capable of achieving this task. In the classroom setting, for example, GenAI could support the conversion of spoken feedback sessions into a structured textual form, freeing researchers from transcription burdens or supporting the classification and categorization of responses from students.

In this article, we report on the use of ChatGPT as a key methodological innovation when analyzing leading academic and educational publishing house policies regarding the use of ChatGPT in authoring research (Perkins & Roe, 2024). The aim of this article is to highlight a specific use case in which ChatGPT and other GenAI tools can help offer greater depth in the analysis of textual data, leading to deeper insights and interpretations. This exploration offers readers an understanding of the methodological nuances, the synthesis of AI-generated outputs with human interpretations, and the challenges and advantages of embedding GenAI tools in qualitative research. Simultaneously, we offer a roadmap for others who wish to conduct educational research using an integrative approach that balances manual analysis with GenAI-assisted analysis and explores the ramifications of using GenAI for qualitative data analysis in the field of education, given the current limitations of these tools.

Through a case study approach, we share our firsthand experiences with GenAI tools, notably ChatGPT, in the context of inductive thematic analysis.

We focus specifically on two questions:

1. How can GenAI tools be used to support qualitative data analysis in educational research?
2. Given the unpredictable nature of GenAI outputs, what implications have emerged for consistent and replicable research if we integrate GenAI tools into methodologies?

### **Case study: Inductive thematic analysis with ChatGPT**

To explore the possibilities of GenAI in supporting qualitative research, we designed a method that integrates traditional inductive thematic analysis with the advanced data processing capabilities of GenAI tools. This choice

was a conscious strategy to obtain a deeper analysis than might otherwise be obtained by a single method alone. We were interested in exploring the differences between the approaches taken in a manual and GenAI-supported methodology while simultaneously cross-checking the validity of each approach.

The methodology employed in this study involved conducting a comprehensive web search to collect and classify policies related to the use of AI tools in academic research from various publishers' websites, which were then combined to form a master list. This was edited to remove any known or suspected instances of publishers of predatory journals and then supplemented with web searches including terms such as 'AI/ChatGPT journal/publisher policies' to identify specific journals, publishers, or publishing-related groups/institutions that had policies related to the use of GenAI tools. This resulted in a master list of 107 entities for review. We then conducted manual searches on the websites of all entities to explore whether a policy was in place relating to the use of Generative AI tools. From this manual search, we identified 36 entities that had an AI policy, eight of which were replicas. Replicas were due to either the use of wording from other entities, such as The Committee of Publication Ethics (COPE), or because the publishing houses were imprints of broader publishing groups and therefore did not have their own policy. This resulted in a final list of 28 unique policies for analysis.

The mode of analysis chosen was inductive thematic analysis (TA) based on the guidelines presented by Braun and Clarke (2006). This particular approach to thematic analysis has become a mainstay in educational research when working with qualitative data, owing to its simplicity, flexibility, and propensity for generating deep insights into different types of data. Furthermore, TA is not theoretically bounded, as is the case with other types of qualitative research frameworks, such as grounded theory.

The methodological innovation that we pioneered in this research project was combining both GenAI and traditional manual analysis of qualitative data using an inductive thematic analysis framework at the semantic level. Epistemologically, this research method was employed from a realist perspective (Braun & Clarke, 2006). As the data collected is related to academic policies, and we wished to explore the nature and framing of the policies, in this case the data was more explicit and 'to the point' than may be found in other domains, for example, in long-form interview data. Consequently, this approach may be applicable only to certain datasets. Given the limitations of current FMs, latent TA may not be suitable for use at the present time with a Gen-AI tool if the researcher expects to analyze more complex forms of meaning expression, such as sarcasm or humor.

To develop this research method, we first decided that both researchers should maintain a separate dataset and not communicate with one another regarding findings until the individual analyses were complete, thereby increasing the objectivity of the results. We followed a step-by-step process for TA, as shown in Table 1.

Table 1: Comparison of traditional versus ChatGPT supported analysis.

Step	Traditional analysis	ChatGPT-assisted analysis
Data familiarization	Collection of data and re-reading of data, noting down initial ideas in word processing software.	
Generating initial codes	Systematic coding of data using traditional software (MaxQDA).	Structuring data into a tabular form, suitable for importing into the GenAI tool.
Searching for themes	Researcher collates codes into prospective themes.	Writing and re-writing prompts to instruct the GenAI tool what is required from the output.
Reviewing themes	Inspecting the themes and reviewing their relation to the codes and dataset.	Inspecting the themes given by the GenAI tools and cross-referencing themes with specific examples from the data for validation purposes.
Defining and naming themes	Continuing analysis and specification of each theme.	Re-running the analysis at different points in time to enhance validity of themes
Combining and Contrasting results	Results are compared by both researchers side by side and areas of congruence are identified in the themes and dataset, final set of themes are produced.	
Producing the report	Creating the scholarly output of the paper.	

### Strengths and weaknesses of the research method

Overall, the method helped to generate two different sets of insight into this data and was productive in answering our research question. We identify several specific strengths associated with this method, namely increased objectivity, efficiency, and additional cognitive support for the researcher. However, the weaknesses of the method include its viability for more complex datasets and the additional time invested in developing necessary GenAI-related skills to produce the desired output.

One of the most interesting aspects of our study is the experience of combining AI-derived insights with manual analysis. Each method brought its strengths to the table. GenAI tools, with their rapid processing capabilities, can sift through vast amounts of data and provide initial themes much more quickly than a traditional analysis, along with a higher number of themes, demonstrating the ability to extract a higher level of granularity from the themes. However, the depth, context, and the subtle nuances were better captured through manual scrutiny. The final synthesis of themes required collaboration, debate, and a deep understanding of both methods. Matching the GenAI themes with those identified in the traditional analysis became a challenging yet revealing process, demonstrating the unique lenses through which humans and AI perceive data.

Engaging with ChatGPT as a research tool, however, has brought its own set of challenges. At times, the tool tended to produce 'hallucinations', generating quotes or data that did not exist in the original dataset. Such anomalies necessitated continuous cross-referencing with our primary data to ensure the integrity of our findings. Furthermore, creating the right prompts for ChatGPT to produce the required output was an iterative and often frustrating process. We realized that obtaining the desired responses from the tool required fine-tuning of our queries. This experience underscored an important facet of using GenAI tools: while they can automate certain processes and offer unique insights, they cannot eliminate the need for human involvement. An additional concern here is one of replicability – GenAI tools are stochastic in nature, and it is unlikely that repeating the prompts used in our analysis



would result in obtaining the same outputs from the tools. The frequent and unclear updates to the underlying GPT4 model further complicate this.

Upon completion of the individual analyses, synthesizing the results was the next challenge. When we compared the insights from the traditional approach with those from the AI-assisted method, we found both overlaps and unique perspectives in the analyses that required significant debate between the researchers regarding which themes were more or less valid for the dataset. This phase involved extensive discussion and collaboration to identify the final set of themes (cf. Perkins & Roe, 2024), and to ensure they were both comprehensive and grounded in the data. It became evident that while the GenAI tool used had strong capabilities in identifying themes for further exploration, a deep understanding of the intricacies and potential pitfalls of using this software was essential to obtain the best possible results and avoid AI hallucinations. This analysis also underscores the importance of human touch in research. While GenAI tools can facilitate, assist, and expedite processes, they cannot replace the unique human ability to interpret, contextualize, and provide depth to findings. Our experience mirrored a broader academic sentiment: GenAI tools are transformative, but they serve best as co-pilots, enhancing and complementing human capabilities rather than replacing them.

### Limitations and future research areas

The advent and integration of GenAI tools in academic research have the potential to bring about a transformation in research methodologies and data analysis. However, the potential of these tools is not without challenges. The stochastic nature of GenAI tools, which can lead to varied outputs even with the same inputs, presents a major challenge for research replicability. Moreover, the rapid pace of technological evolution and updates to these tools may risk making certain methodologies or analyses obsolete. This was seen throughout the course of the research being discussed, in which the specialist ChatGPT mode of Code Interpreter/Advanced Data Analysis was folded into the overall ChatGPT Plus package. These rapid changes can complicate longitudinal studies or any attempts to compare new results with past data. Additionally, while these tools promise efficiency and depth, there is a growing concern about over-reliance on AI by both researchers and students, which might eclipse the essential human insight in the research process.

In terms of future research directions, there's a pressing need for comparative studies. Such studies should aim to discern the efficacy of GenAI tools against traditional research methods in various academic fields. By understanding how the benefits of these tools can enhance the research process, the academic community can harness their strengths more effectively while simultaneously limiting the potential negative impacts of their challenges and weaknesses. Furthermore, as the role of GenAI in research becomes more pronounced, a deeper consideration of the ethical considerations surrounding these tools cannot be avoided. Issues related to data privacy, potential AI-generated

biases, and the broader societal implications of AI-driven research warrant further exploration and debate. Given the delays demonstrated in integrating the considerations of GenAI into academic integrity policies (Perkins & Roe, 2023), it is important that clear policies and guidelines on how to effectively and ethically integrate GenAI tools into research activities are provided to both faculty and student researchers.

Although the current research focuses on qualitative analysis, a further area for investigation is how GenAI tools can be integrated into quantitative research methodologies. With their ability to simplify complex data tasks, GenAI tools, especially in multimodal form, can produce visual analytics to support pattern identification. Pairing GenAI with Python libraries that enable more advanced quantitative analysis techniques also democratizes the research process, helping researchers without extensive technical backgrounds to engage in advanced analyses. For instance, tasks such as regression analysis, previously reserved for those with specific expertise, can now be approached through simple language prompts to GenAI tools.

### Conclusion

The integration of GenAI tools into the academic realm signifies more than just technological advancement; it embodies a true paradigm shift in how research is conceptualized and executed. These tools can act as co-pilots: augmenting the capabilities of human researchers rather than seeking to replace them. With the assistance of GenAI tools, researchers can explore large volumes of data using natural language prompts, and without specialized software, therefore democratizing the research process and deriving insights that might have remained obscured in traditional methodologies. Although GenAI tools can support academic research, they do not yet have the ability to replace it. The same can be said for educational processes. Research has shown that in tasks such as generating educationally aligned assessment items, the technology fails to match human performance (Khademi, 2023); therefore, a high degree of criticality needs to be applied when planning to use these tools in uncharted territory. While they bring new benefits to research processes, it is the human researcher's intuition, expertise, and interpretative skills that breathe life into data, transforming it into meaningful knowledge. As we move forward, we should consider this a form of symbiosis, with GenAI tools and human researchers collaborating to amplify the other's strengths, thereby pushing forward what would otherwise be possible in academic research.

In summary, our case study provides a microcosm of the larger challenges inherent in integrating GenAI tools into academic research. The challenges we face, from AI hallucinations to theme synthesis, are key identifiers that, as yet these tools are not ready to stand by themselves in research workflows. However, the possible benefits, from rapid data processing to diverse insights, demonstrate their potential as helpful co-pilots for enhancing the effectiveness and efficiency of research results. The key lies in understanding, adapting, and striking the right balance between human expertise and AI capabilities.

## AI usage disclaimer

This study used Generative AI tools to analyze data, create preliminary themes, produce draft text, and revise wording throughout the production of the manuscript and accompanying research note. Multiple modes of ChatGPT over different time periods were used, with all modes using the underlying GPT-4 Large Language Model. The authors reviewed, edited, and take responsibility for all outputs of the tools used in this study.

## Data availability statement

The underlying data that is referred to in this research note has been published and is publicly available on Figshare at the following link <https://doi.org/10.6084/m9.figshare.24124860.v1>.

This project contains the following underlying data:

- Gen AI policies Academic Publisher .xlsx. (All data related to AI policies including full policies as downloaded and URLs).

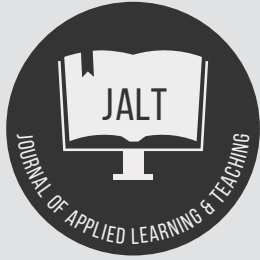
The data is available under the terms of the CC-BY 4.0 license.

## References

- Abd-Elaal, E.-S., Gamage, S. H., & Mills, J. E. (2022). Assisting academics to identify computer generated writing. *European Journal of Engineering Education*, 1–21. <https://doi.org/10.1080/03043797.2022.2046709>
- Anderson, N., Belavy, D. L., Perle, S. M., Hendricks, S., Hespanhol, L., Verhagen, E., & Memon, A. R. (2023). AI did not write this manuscript, or did it? Can we trick the AI text detector into generated texts? The potential future of ChatGPT and AI in sports & exercise medicine manuscript generation. *BMJ Open Sport & Exercise Medicine*, 9(1), e001568. <https://doi.org/10.1136/bmjsem-2023-001568>
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). *Education in the era of generative Artificial Intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning* (SSRN Scholarly Paper 4337484). <https://doi.org/10.2139/ssrn.4337484>
- Bowman, S. R. (2023). *Eight things to know about large language models*. <https://doi.org/10.48550/arXiv.2304.00612>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Caruccio, L., Cirillo, S., Polese, G., Solimando, G., Sundaramurthy, S., & Tortora, G. (2024). Can ChatGPT provide intelligent diagnoses? A comparative study between predictive models and ChatGPT to define a new medical diagnostic bot. *Expert Systems with Applications*, 235, 121186. <https://doi.org/10.1016/j.eswa.2023.121186>
- Chaka, C. (2023). Detecting AI content in responses generated by ChatGPT, YouChat, and Chatsonic: The case of five AI content detection tools. *Journal of Applied Learning and Teaching*, 6(2), 94–104. <https://doi.org/10.37074/jalt.2023.6.2.12>
- Chen, T.-J. (2023). ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association*, 86(4), 351–353. [https://journals.lww.com/jcma/\\_layouts/15/oaks.journals/downloadpdf.aspx?an=02118582-202304000-00001](https://journals.lww.com/jcma/_layouts/15/oaks.journals/downloadpdf.aspx?an=02118582-202304000-00001)
- Chopra, F., & Haaland, I. (2023). *Conducting Qualitative Interviews with AI* (SSRN Scholarly Paper 4583756). <https://doi.org/10.2139/ssrn.4583756>
- Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1–12. <https://doi.org/10.1080/14703297.2023.2190148>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Elali, F. R., & Rachid, L. N. (2023). AI-generated research paper fabrication and plagiarism in the scientific community. *Patterns*, 4(3), 100706. <https://doi.org/10.1016/j.patter.2023.100706>
- Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching*, 6(1), 57–63. <https://doi.org/10.37074/jalt.2023.6.1.22>
- Foltynek, T., Bjelobaba, S., Glendinning, I., Khan, Z. R., Santos, R., Pavletic, P., & Kravjar, J. (2023). ENAI Recommendations on the ethical use of Artificial Intelligence in Education. *International Journal for Educational Integrity*, 19, 12. <https://doi.org/10.1007/s40979-023-00133-4>
- Haman, M., & Školník, M. (2023). Using ChatGPT to conduct a literature review. *Accountability in Research*, 0(0), 1–3. <https://doi.org/10.1080/08989621.2023.2185514>
- Hassoulas, A., Powell, N., Roberts, L., Umla-Runge, K., Gray, L., & Coffey, M. J. (2023). Investigating marker accuracy in differentiating between university scripts written by students and those produced using ChatGPT. *Journal of Applied Learning and Teaching*, 6(2), 71–77. <https://doi.org/10.37074/jalt.2023.6.2.13>
- Kasneji, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A.,

- Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Khademi, A. (2023). Can ChatGPT and Bard generate aligned assessment items? A reliability analysis against human performance. *Journal of Applied Learning and Teaching, 6*(1), 75-80. <https://doi.org/10.37074/jalt.2023.6.1.28>
- Köbis, N., & Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior, 114*, 106553. <https://doi.org/10.1016/j.chb.2020.106553>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), 100790. <https://doi.org/10.1016/j.ijme.2023.100790>
- Limna, P., Kraiwani, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching, 6*(1), 75-80. <https://doi.org/10.37074/jalt.2023.6.1.32>
- Lundgren, M., Squatrito, T., & Tallberg, J. (2018). Stability and change in international policy-making: A punctuated equilibrium approach. *The Review of International Organizations, 13*(4), 547–572. <https://doi.org/10.1007/s11558-017-9288-x>
- Mills, A., Bali, M., & Eaton, L. (2023). How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *Journal of Applied Learning and Teaching, 6*(1), 16-30. <https://doi.org/10.37074/jalt.2023.6.1.34>
- OpenAI. (2022, November 30). *ChatGPT: Optimizing language models for dialogue*. OpenAI. <https://openai.com/blog/chatgpt/>
- OpenAI. (2023, August 28). *Introducing ChatGPT enterprise*. <https://openai.com/blog/introducing-chatgpt-enterprise>
- Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching & Learning Practice, 20*(2). <https://doi.org/10.53761/1.20.02.07>
- Perkins, M., Furze, L., Roe, J., & MacVaugh, J. (2023). *Navigating the generative AI era: Introducing the AI assessment scale for ethical GenAI assessment* (arXiv:2312.07086). arXiv. <https://doi.org/10.48550/arXiv.2312.07086>
- Perkins, M., & Roe, J. (2023). Decoding academic integrity policies: A corpus linguistics investigation of ai and other technological threats. *Higher Education Policy*. <https://doi.org/10.1057/s41307-023-00323-2>
- Perkins, M., & Roe, J. (2024). *Academic publisher guidelines on AI usage: A ChatGPT supported thematic analysis* (12:1398). F1000Research. <https://doi.org/10.12688/f1000research.142411.2>
- Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2023). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics*. <https://doi.org/10.1007/s10805-023-09492-6>
- Pichai, S. (2023). *Google I/O 2023: Making AI more helpful for everyone*. <https://blog.google/technology/ai/google-io-2023-keynote-sundar-pichai/#ai-products>
- Rahimi, F., & Abadi, A. T. B. (2023). ChatGPT and publication ethics. *Archives of Medical Research, 54*(3), 272–274. <https://doi.org/10.1016/j.arcmed.2023.03.004>
- Roe, J., Renandya, W., & Jacobs, G. (2023). A review of AI-powered writing tools and their implications for academic integrity in the language classroom. *Journal of English and Applied Linguistics, 2*(1). <https://doi.org/10.59588/2961-3094.1035>
- Salah, M., Al Halbusi, H., & Abdelfattah, F. (2023). May the force of text data analysis be with you: Unleashing the power of generative AI for social psychology research. *Computers in Human Behavior: Artificial Humans, 1*(2), 100006. <https://doi.org/10.1016/j.chbah.2023.100006>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching, 6*(1), 31-40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guerrero-Dib, J., Popoola, O., Šigut, P., & Waddington, L. (2023). Testing of detection tools for AI-generated text. *International Journal for Educational Integrity, 19*(1), Article 26. <https://doi.org/10.1007/s40979-023-00146-z>
- Xames, M. D., & Shefa, J. (2023). ChatGPT for research and publication: Opportunities and challenges. *Journal of Applied Learning and Teaching, 6*(1), 390-395. <https://doi.org/10.37074/jalt.2023.6.1.20>
- Zhai, X. (2022). *ChatGPT user experience: Implications for education* (SSRN Scholarly Paper 4312418). <https://doi.org/10.2139/ssrn.4312418>

Copyright: © 2024. Mike Perkins and Jasper Roe. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Reading between the lines: The necessity of books

Sam Choon Yin<sup>A</sup>

A

*Dean (Academic Partnerships), Kaplan Higher Education Academy*

---

### Keywords

Books;  
educational technology;  
higher education;  
knowledge;  
reading.

---

### Abstract

This opinion piece delves into the unwavering significance of books as an enduring and robust medium of education. It traces the historical trajectory of books, from ancient scrolls to contemporary printed volumes, elucidating their consistent role in education and knowledge dissemination. Books serve as indispensable repositories of human wisdom, culture, and progress, preserving the intellectual heritage of bygone eras and steering society toward enlightenment. Amidst the digital age's proliferation of screens and electronic devices, this paper argues that books in the physical form maintain their charm, captivating readers and learners alike. Their tangible presence and tactile engagement provide a unique haven for knowledge seekers, serving as steadfast companions in the lifelong pursuit of learning. In an era marked by rapid technological advancement, books stand as a testament to the enduring worth of written knowledge. As we embrace innovation, it is crucial to acknowledge the timeless wisdom contained within books' pages, forever prepared to guide us on our journey to comprehension and enlightenment.

---

### Correspondence

choonyin.sam@kaplan.com <sup>A</sup>

---

### Article Info

Received 6 November 2023

Received in revised form 29 November 2023

Accepted 3 January 2024

Available online 5 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.2>

If you want to be a writer, you must do two things above all others; read a lot and write a lot. There's no way around these two things that I'm aware of, no shortcut (King, 2012, p. 164).

## Introduction

Unlike food, human beings do not need to read to survive. The question arises: Why do we read? Book readers would argue that delving into the pages of a book propels us into the imaginative worlds created by others while simultaneously cultivating our unique perspectives. In this sense, reading nurtures our imagination, exposing us to a spectrum of experiences far removed from our personal lives.

Indeed, reading bolsters our cognitive abilities. As we immerse ourselves in diverse literary landscapes, our academic and professional capabilities flourish. We begin to challenge established norms, decipher news with greater insight, and communicate the implications of current events to others. As a result, dictators have historically sought to suppress reading and ban books as they understand their potential to empower and embolden individuals (Ferner & Meyns, 2021). This is also why, for years, slaves were forbidden to learn to read, and girls in some nations have suffered violence in response to their pursuit of education (Cornelius, 1991; GCPEA, 2019; Whitehead, 2016). The gift of education and books shines a beacon of hope, especially to those facing adversity, illuminating a path toward a brighter future.

This paper focuses on the tangible form of books as the medium for reading. These books, consisting of collated sheets bound along the spine and encased in protective covers, facilitate multifaceted engagements with their content. In an age dominated by digital screens and electronic devices, we contend that physical books continue to exert their appeal, enchanting readers and learners alike. Their tangible presence and tactile engagement offer a unique sanctuary for those seeking knowledge, serving as unwavering companions in the lifelong quest for learning. In a time of rapid technological progress, books stand as a testament to the lasting value of written knowledge. As we welcome innovation, it is paramount to recognise the timeless wisdom encapsulated within the pages of books, always ready to lead us on our path to understanding and enlightenment.

I approach the subject by studying the historical context of books to reveal how they have adapted and coexisted with emerging technologies. Exploring the historical trajectory of books and writing provides invaluable insights into their enduring significance in our contemporary world. A historical perspective also unveils the evolutionary journey of the written word, illuminating its profound impact on human culture, communication, and knowledge dissemination. This perspective underscores the unique qualities of physical books, such as the sensory experience of holding a printed page, which digital formats cannot fully replicate. Amidst the digital revolution, the question lingers: What is the role of physical books in the contemporary world? This exploration leads us to conclude that, even in the digital age, books in their traditional form retain significant relevance. To read

one's work, the work has to be written. It makes sense to begin our story with writing.

## Writing

Writing was likely pioneered by the Sumerians in Uruk, although it also emerged independently around five thousand years ago in Egypt and China. Uruk was a major urban and cultural hub of Sumerian civilisation (Graeber & Wengrow, 2022). The city played a significant role in shaping Sumerian culture, trade, governance, and religious practices. It is likely that cuneiform script began to take shape in Uruk around 3300 BCE, with early manifestations evident in numerical tablets and administrative notations. Initially, the primary function of writing in the city's temples was bookkeeping. Over the course of millennia, Uruk's temples became the very sites where cuneiform script eventually became obsolete, having evolved to document a wide range of matters, including the earliest recorded literature and legal codes in the world (see Rowe & Levine, 2006; Fischer, 2021).

The Sumerians and, later, the Egyptians used rudimentary symbols to convey basic ideas and concepts. However, as societies grew more complex, the need for a more efficient and simpler writing system became evident. Around 2000 BCE, the Phoenicians, a seafaring people from the eastern Mediterranean, introduced a groundbreaking innovation – the alphabet (see Vallejo, 2022). Unlike the intricate hieroglyphs of the Egyptians or the cuneiform script of the Sumerians, the Phoenician alphabet consisted of just 22 characters, each representing a consonant sound. This compact system allowed for greater speed and flexibility in writing, making it suitable for various languages. Alphabets were used extensively by traders to keep track of what was bought and sold.

The Phoenician alphabet served as the foundation for numerous other scripts. As documented in Irene Vallejo's (2022) *Papyrus*, the Greeks adopted and adapted it, adding vowels and refining the characters to suit their language. This Greek alphabet, in turn, passed on to the Romans. The Latin alphabet spread across the Roman Empire and eventually became the basis for many modern European languages. Writing meant that rulers could create legal codes and religious leaders could spread ideas through religious texts. The Arabic alphabet, derived from ancient scripts, was widely adopted across the Islamic world (Goody, 2006).

Meanwhile, other civilisations were developing their own writing systems. In India, the Brahmi script emerged around the 3rd century BCE, paving the way for various scripts across the Indian subcontinent, including Devanagari and Tamil. In East Asia, Chinese characters evolved over centuries, with each character representing a syllable or a concept. These characters formed the basis for Japanese Kanji and Korean Hanja (see Ferner & Meyns, 2021; Fischer, 2021).

The invention of writing has been criticised by various scholars. Socrates, for instance, warned that writing could lead to ignorance and, ultimately, death of memory as we

no longer feel responsible for remembering knowledge. Writing was also seen as a means to convey information without truly imparting knowledge, giving the illusion of wisdom without necessarily imparting true understanding. Indeed, in ancient times, oral history held a superior level of reliability, allowing ideas to adapt and evolve naturally, as they would during spoken exchanges. Throughout much of history, even well into the 11th century, confidence in written records did not manifest immediately or unconditionally (Cohen, 2023). When individuals sought information about the past, they did not turn to books or scrolls. Instead, they relied on the wisdom passed down through generations, shared by their elders.

But clearly without writing, we would wake up each morning with fading memories of yesterday and no way to plan for tomorrow, let alone the day after tomorrow. The advent of writing liberates us from the constraints of memory, transforming the repository of knowledge into a tangible archive capable of limitless expansion. In addition, the transition from spoken to written language helps to crystallise the ephemeral nature of one's identity into a concrete and lasting manifestation. Through the act of writing, words cease to be transient vocalisations and solidify into a tangible representation, thereby extending the longevity and impact of the speaker's identity beyond the temporal confines of verbal communication. Writing, being more enduring and constant than speech, ensures that the legacy of heroes endures. Their feats continue to live on through written accounts, ensuring their presence in our consciousness to this day.

Unlike oral traditions, which tend to favour familiar forms and established ideas for the sake of audience recognition, the written word beckons us towards uncharted intellectual horizons. Intellectuals play a crucial role in materialising thoughts and ideas by expressing them through writing, books, and various forms of records and documents. This transformation from the intangible thoughts and ideas to the tangible facilitates their continued existence, and enables ideas and concepts to evolve over time. In the absence of such tangible forms, thoughts and ideas risk fading into oblivion, losing their reality as if they had never existed. With writing, readers are also afforded the luxury of time, permitting contemplation and introspection upon novel concepts. We can visually behold words and ponder them at our own pace, free from the rushing current of spoken discourse.

## Books

Early evidence of writing in the form of pictorial images and paintings emerged in the caves of Lascaux in southwest France, dating about 15,000 BCE and Cueva de las Monedas in Spain in the Ice Age (Lyons, 2013). Cuneiform, an ancient writing system, emerged around the close of the 4th millennium BCE. It involved inscribing signs and numbers onto clay tablets with a pointed stylus. These tablets, comparable in size to modern credit cards, were then left to dry in the sun. This method of recording, as described by Lyons (2013), was prevalent in Mesopotamia and employed for documenting taxation and legal affairs.

The transition of ancient Egyptians from prehistory to history was marked by the innovation of a medium beyond stone, metal, or leaves for inscription. They ingeniously harnessed the potential of the papyrus plant stem, abundant in the Nile delta, which previously served various purposes like crafting furniture, baskets, ropes, and boats (Vallejo, 2022). The Greeks referred to this plant as *biblos*, a term that eventually evolved into the English word 'book,' illustrating the profound impact of this innovation on the evolution of written communication. Papyrus was used to record signs and numbers in Egypt, Greece, Rome and all over the Mediterranean world. Plato, Thucydides and Cicero all wrote on papyrus (Lyons, 2013).

In the year 105 CE, Cai Lun, a eunuch of the imperial court in China, revolutionised the world by inventing paper. His method, which involved using old rags, hemp, tree bark, and fishing nets, laid the foundation for the papermaking process still employed today (Smith, 2022). This innovative art of papermaking, originating from China and subsequently adopted by the Arabs, spread across the Islamic world during the first millennium.

A major revolution of the book was the invention of the codex, which originated in the Christian world of the 2nd and 3rd centuries (Duncan, 2022). Books became a collection of separate sheets loosely attached or sewn to each other. The key advantage of parchment codex is that writers were able to write on both sides of the sheet. Parchment codex had a sturdy individuality, allowing for easy storage. Concurrently, the concept of an index emerged. The index is an invention of the codex era, serving as a time-efficient tool akin to a map that mediates the relationship between authors and their readers. Its role is significant, offering a navigational aid that enhances the accessibility and utility of written works. Reading a book from cover to cover is an investment of hours, something we have to make time for. With the aid of a good index, checking a reference is the work of seconds. Notably, the concept of the index parallels the modern-day concept of a webpage index. In the contemporary digital landscape, when we conduct a Google search, we are not directly exploring the entirety of the internet. Rather, as Duncan (2022) told us, we are navigating Google's curated index of web content.

Prior to the invention of the movable-type printing press, book production was cost-prohibitive due to sheet-based manufacturing. Consequently, books were predominantly utilised by scholars and for institutional purposes by civil or military officials. In China, beginning with Confucius (551-479 BCE), books primarily functioned as educational tools, serving as vessels of knowledge that spanned philosophy, medicine, astronomy, and cartography. In the Western context, under the rule of Charlemagne (742-814), the inaugural emperor of the West after the fall of Rome in the era of Christian rule, monasteries were assigned the responsibility of producing and disseminating books extensively to promote scholarly endeavours and intellectual development (Pettegree & Der Weduwen, 2021). Apart from the individuals of privilege, the practice of book collecting was primarily confined to the affluent, for whom acquiring and possessing books served more as a means of showcasing wealth and status rather than a pursuit of reading (Vallejo, 2022). This means that the

collector could not afford his books to look inexpensive. He had to invest in the finest copies, decorated with his coat of arms and numerous illustrations, if only to impress upon his visitors the patron's respect for learning and love of books. Amassing books was a privileged undertaking.

From the 12th century onwards, the monasteries in the West gradually lost their hegemony over the production and circulation of books. Serious challenges emerged from new institutions, cathedrals, schools and universities, which grew in towns in France, Germany and England. Growth of courts and universities stimulated literary activity and further intensified the need for books. Niccolo de Niccoli (Italian Renaissance; 1364 – 1437), a book hunter, founded the cursive writing type to allow the pen to move more freely, creating greater efficiency in copying, and new markets including a new class of book dealers, the *cartolai*, as middlemen between scribes and the clients, and traders of books (Pettegree & Der Weduwen, 2021).

New technologies appeared in rapid succession. Johannes Gutenberg, an inventor hailed for his revolutionary contribution, is credited with introducing movable type to the world. This innovation enabled the arrangement of individual letters into coherent forms that could be reproduced numerous times through the application of a printing press. Born around 1400 in Mainz, located along the River Rhine, Gutenberg's background as a goldsmith significantly influenced the evolution of printing. The precision required in crafting fine metal pieces played an instrumental role in creating accurate print impressions. The advent of the printing press marked a monumental shift in the pace of book production. Notably, the Gutenberg Bible, spanning an impressive 1,282 pages, made its debut in 1455 (Smith, 2022). While a significant portion of this Bible was printed on paper, certain sections were rendered on materials associated with traditional manuscript work, such as vellum or calfskin.

Johannes Gutenberg's movable type technology is particularly well-suited for languages that utilised a limited set of alphabetic symbols (Simonds, 2017; Bickerts, 2006). This ingenious method allowed for the arrangement and rearrangement of these symbols to create any desired word. Additionally, he introduced an oil-based ink, a significant improvement over the water-based ink utilised in the manual copying of texts.

While printing had already been pioneered in China centuries earlier, the Chinese approach required the carving of numerous individual characters onto separate wooden blocks. In contrast, Gutenberg's movable type allowed for the reuse of a relatively small set of letters, which could be combined in various ways to create different texts. The impact of Gutenberg's innovation on book production cannot be overstated. Setting up his printing press in the city of Mainz, Gutenberg triggered a transformative shift. Within less than half a century of his invention, nearly 300 cities across Europe had embraced printing presses, leading to the production of an astonishing twenty million copies of various books (Grayling, 2022). This rapid expansion and replication of texts marked a remarkable leap in the dissemination of knowledge and information throughout the

continent, revolutionising the way information was shared and accessed by people from all walks of life. Innovation in production and distribution, the steam press and the railways, multiplied the quantity of books available and helped bring printed materials to consumers. The sheer quantity of books in circulation drove prices down (Pettegree & Der Weduwen, 2021). The proliferation of books also heralded the inception of spectacles, a technological marvel that brought distant objects astonishingly close. This innovation revolutionised visual perception, enabling people to perceive faraway entities as if they were within arm's reach. Moreover, this era witnessed the dawn of the chapter system, a structural innovation that divided books into smaller, manageable segments (Duncan & Smyth, 2019). This division facilitated uninterrupted reading, allowing individuals to immerse themselves in a book for an extended period and pause at designated intervals for reflection and contemplation.

### **The emergence of books as a potent weapon**

Books became a powerful medium to convey and spread ideas. The act of contemplating the world and engaging in profound thought owes its existence to the existence of books, writing, and the act of reading itself. The Protestant Reformation encouraged the spread of vernacular languages in print, and the first major printed text to be translated from Latin to vernacular languages was the Bible (Lyons, 2013). Protestant leaders believed that the message of the Bible should be accessible to all Christians in their own languages – Martin Luther's ideal of the "priesthood of all believers", a world in which ordinary people could consult the word of God for themselves, without the guidance and interpretation of the clergy. Luther's "Ninety-Five Theses" would not have spread so far and wide without a printing press to publish his posters. When Luther's "Ninety-Five Theses" emerged, as noted by Brian Cummings (2022), only a few thousand individuals encountered a copy of the broadsheet. Nevertheless, authorities apprehended the challenge of eradicating it, underscoring the pervasiveness of writing and printed books.

President Franklin D. Roosevelt eloquently emphasised the influential role of books as tools of empowerment, stating that they wielded a potency akin to formidable weaponry. He likened books to battleships, attributing them with the sturdiest armour, the longest operational range, and the most formidable artillery (Smith, 2022). In a concerted effort, the American Council of Books in Wartime authored over 1,300 titles intended to fortify the intellectual and emotional resolve of the American populace with enduring tools of enlightenment and strength. These volumes were distributed in paperback format, conveniently sized to fit within a soldier's uniform pocket. As highlighted in Smith (2022), the books served multifaceted purposes: nurturing determination, unveiling the true nature of the adversary, disseminating technical knowledge relevant to training and combat, providing a source of relaxation, instilling inspiration to bolster morale, and elucidating the war's objectives. Throughout this period, a diverse range of literary genres was employed as ammunition, spanning poetry, mystery novels, serious non-fiction, humour, and contemporary bestsellers. The sentiments of soldiers further underscored

the importance of books, as they were received with the same welcome as letters from home and even revered akin to pin-up images.

The utilisation of paperback books during wartime not only popularised innovative publishing practices but also galvanised advancements in publishing, including the production of affordable books. Visionary publishers such as Allen Lane of Penguin Books and Robert de Graff of Pocket Books played instrumental roles, particularly in the 1930s. Their efforts not only challenged entrenched copyrights and distribution monopolies but also set the stage for broader transformations in the publishing industry (Smith, 2022).

As books got cheaper, more people were able to afford to buy their own books and build their own collections. The book industry in the 1960s experienced a boost with the growing influence of television and cinemas. Novels began to be adapted into movies, and writers responded by producing biographies of movie stars. Universities supported the growth of highly specialised fields of knowledge to keep pace with the incoming data, leading to the growth of the number of skilled professionals (Rumsey, 2016). The sheer growth of book titles and information in general led to demand for information management infrastructure – libraries, museums and skilled staff to manage the assets. More than ever, there was a need for libraries and librarians to help readers find what they needed. In the United States, Andrew Carnegie founded the modern American public library, offering a space for individuals to seek entertainment and education. Rumsey (2016) reported that in 1836, the Library of Congress in the United States housed 24,000 volumes, four times more than it had 20 years earlier. The British Museum (now the British Library) had 180,000 titles.

Books, like writing, have been a subject of criticism. History, as the saying goes, is written by the winners, hence creating biased stories and narratives. The fact is that authors write books about almost everything, even if they could not yet be completely sure of what it was. "The Theory of Everything" as the legendary Charles Handy (1995, p. 17) wrote, "is a fallacy in a possibility of perfection." Even esteemed writers can unconsciously overlook the pursuit of objective history, instead opting to include what aligns with the particular agenda they have chosen to follow. In their narratives, they reveal truths that were perceived through their own lenses, reflecting their individual experiences. These books are often birthed from a myriad of emotions – despair, anger, the yearning for revenge, the pursuit of power, or a sense of personal calling that guides them through each chapter. Edward Said (1994) aptly coined this literary perspective "Orientalism", highlighting the complex interplay between history, identity, oppression, and personal motivation.

Writing and books are like chronicles of our history, keeping alive memories of times, a wellspring of hope, solace, and empathy. Our roles as engaged citizens, whether during peaceful periods or in times of turmoil, can be seen as an extension of our reading. In this regard, books hold the power to inculcate values and transform behaviour through storytelling and knowledge dissemination. By presenting characters facing moral dilemmas and showcasing their choices and consequences, books provide readers with

vicarious experiences that prompt reflection on their own values and decisions. They offer insights into diverse perspectives, fostering empathy and understanding. Furthermore, books can present well-researched arguments and evidence, influencing readers' beliefs and attitudes. Over time, the accumulation of knowledge and exposure to different ideas can lead to profound shifts in individual behaviour, as readers internalise new values and perspectives encountered in their literary journeys.

On that note, the fear of books inculcating wrong values has resulted in the burning of books. Qin Shi Huang (259-201 BCE), the emperor of the Qin dynasty (reigned 221-210 BCE) in northwestern China, orchestrated the burning of books as part of his efforts to shape the ideology of the newly unified Chinese empire. This decree, which occurred around 221 BCE, involved the destruction of all books not related to agriculture, medicine, or prognostication. Notably, historical records of the Qin state and books housed in the imperial library were spared from the flames (Ferner & Meyns, 2021). In the evening of 10 May 1933, book burning took place in 34 university towns across Germany organised by Nazi student groups as the climax of their campaign of "Action Against the Un-German Spirit". Thousands of books were transported and piled up for destruction. Works by more than 50,000 authors whom the German leaders considered degenerate were gathered for burning, including works by Einstein, Freud, Marx, Kafka, and Hemingway. Copies of the Hebrew scriptures were also burned. Eleanor Roosevelt, in a newspaper piece "My Day", published on 11 May 1943, cited freedom of speech and thought as crucial to democracy and noted that book burning had the opposite effect; banned authors' "contributions to the thinking of the world are probably greater than they would have been without Hitler's effort at suppression" (cited in Smith, 2022, p. 147). Book burning has become the most highly charged and visible form of attacks on culture.

Despite the United States' assertion of unrestricted reading rights as an embodiment of press freedom and liberty, the reality stands in stark contrast to this claim (see McAllister & Harati, 2023). There is the element of fear, fearful of misinterpretation of everything read in unfamiliar ways, leading to 'undesirable' outcomes. Instances of book bans have emerged, as succinctly highlighted by Smith (2022). Theodore Dreiser's "An American tragedy" fell victim to censorship in Boston, as the Boston District Attorney sought to counter books deemed perilous to the morals of youth. Likewise, Ernest Hemingway's "The sun also rises" encountered censorship in Boston during 1930 due to its language, profanity, and unflinching exploration of themes such as sex, promiscuity, and societal decadence. John Steinbeck's "The grapes of wrath" faced the burning flames of opposition in 1930s California, ignited by a group of farmers resisting labour law reform. The novel was not only incinerated but also exiled from libraries in Illinois, Kansas, and New York State. A more sinister figure, Adolf Hitler, authored "Mein Kampf" during his prison stint in Munich in the early 1920s. The book, a two-volume autobiographical manifesto dedicated to the casualties of the failed November 1923 Nazi party coup, was published in 1925 and 1926. However, it was not until 1933, when Hitler ascended to power, that it morphed into a single-volume bestseller. In



that same year, an abridged American edition translated by E. T. S. Dugdale stirred widespread outrage. The publisher, Houghton Mifflin, faced vehement criticism for providing a platform for Hitler's ideology. This backlash culminated in a petition that condemned the publisher for propagating Hitler's propaganda. The repercussions were tangible, as Houghton Mifflin lost its public contract for textbooks in New York schools.

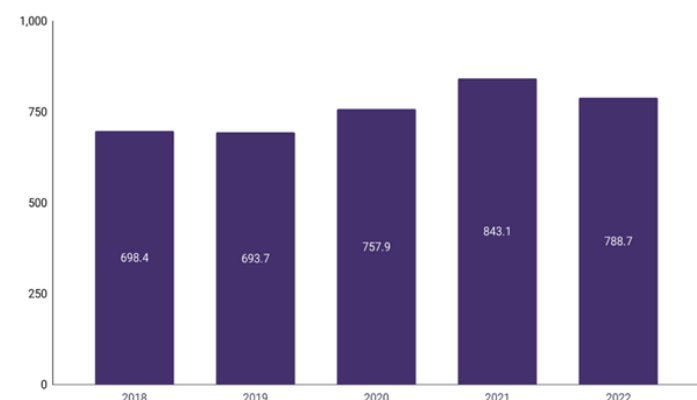


Figure 1: Print book sales in the United States (million units sold). Source: McLoughlin (2023).

While books have faced criticism and burning, printed books persist and thrive. Peter Drucker (1999) distinguished four Information Revolutions, with the second one sparked by the advent of written books, which significantly expanded the dissemination of knowledge, enabling independent information access and exploration, thereby nurturing a culture of learning and intellectual advancement. Accurately tallying the number of printed books proves challenging due to diverse publishing channels, the surge in self-publishing, global variations in practices, the influence of digital formats, incomplete reporting, and the industry's dynamic nature. Nonetheless, there are some indications available. Google calculated that by 2010, 129,864,800 books had been published since Gutenberg and Google wanted to digitise them all (Pettergree & Der Weduwen, 2021). Google does not consider books published post-2010, and it also omits self-published titles. This results in the omission of a substantial number of books from the records. Zaid Gabriel (2004) observed that more than 50 million books had been published. Danny McLoughlin (2023) documented that in 2022, the United States witnessed the sale of over 788.7 million printed book copies. This figure ranked as the second-highest sales record for printed books in the 21st century, coming slightly behind the record-breaking sales of 843 million copies in 2021 (Figure 1). Adult nonfiction remains the dominant category, contributing significantly with over 289 million copies sold, making up 37.8% of all print book sales. Following closely is the adult fiction category, securing the second spot with 189 million copies sold in the same year, constituting 24.7% of print book sales.

## Digital reading revolution: challenges and implications for teaching and learning

The rise in the number of books can be partially attributed to the emergence of e-books, or electronic books. While e-books existed in the early 1970s, it was not until the late 1990s and early 2000s that e-books started to gain significant popularity and become more widely available to the general public. One of the contributing factors was the introduction of dedicated e-readers, such as the Amazon Kindle, which made it more convenient for readers to access digital books. These devices, with their E-ink displays and long battery life, helped popularise e-books and changed the way people read. Since then, e-books have become an integral part of the publishing industry, offering readers a convenient and portable way to access a wide range of literature and other written materials.

E-books mirror their printed counterparts with the number of words arranged on each line. Texts are displayed vertically on e-readers, which are designed in a portrait format. To navigate through the content sequentially, pages are flipped from right to left, and users can conveniently bookmark or underline specific passages. E-books also allow readers to carry an entire library on portable devices, breaking down geographical barriers and enabling instant access to a wide range of content. Additionally, e-books offer features like adjustable font sizes, interactive elements, and search functionalities that enhance the reading experience. The rise of e-books has democratised publishing, enabled self-publishing and increased the availability of diverse voices.

Futurists, like Ray Kurzweil, predicted that with e-books, the printed books would be rarely used (Kurzweil, 1999). It is worth noting that predictions about the decline of books had circulated long before this point in time. In the 19th century, there was a prevailing belief that the advent of daily newspapers would signal the decline of books (Carr, 2016). The argument was that books could not compete with the immediacy of morning and evening papers. Similarly, in the 1950s, as television gained popularity in the Western world, many believed it would spell the end of printed books. With the emergence of computers, printed books are considered a passive medium, lacking the interactions of websites and apps and all other online activities we are accustomed to. There was a widespread expectation of a digital revolution in book publishing. Indeed, as told by Nicholas Carr (2016), at the 2011 Edinburgh International Book Festival, Scottish novelist Ewan Morrison predicted that the digital revolution would lead to the extinction of paper books within the next 25 years.

However, even as new forms of communication and information sharing have emerged, the codex, in the form of printed books that we are familiar with today, has managed to stand the test of time. Nicholas Carr (2016, p. 88) wrote that, in the United States, "e-book sales, which skyrocketed after the launch of Amazon's Kindle in late 2007, have fallen back to earth in recent months, and sales of physical books have remained surprisingly resilient. Printed books still account for about three-quarters of overall book sales in the United States, and if sales of used books, which have been booming, are taken into account, that percentage

probably rises even higher" (Carr, 2016, p. 88). In his book, "Book wars," John Thomson (2021) highlights the ascent of e-books in the American market, which began in 2008-2009 following the introduction of the Kindle. Sales steadily climbed until 2012, at which point they plateaued before experiencing a decline. In contrast, printed books maintained their dominance, consistently representing the majority of total sales. From 2015 onward, printed books accounted for between 80% and 85% of total sales. Drawing on the Pew Research Center survey conducted from 25 January 2020 to 8 February 2021, Michelle Faverio and Andrew Perrin (2022) reported that Americans exhibit a continued affinity for print books, with 65% of adults having read a print book in the past year, making it the most popular reading format. While 30% have explored e-books in the same timeframe, print books maintain their dominance. Interestingly, a majority of readers (33%) engage with both print and digital formats, highlighting a balanced approach. 32% exclusively favour print books, underscoring the enduring appeal of traditional reading. A mere 9% restrict their reading to digital formats, omitting print entirely.

There are two main reasons for this. First, storing facts in digital devices and personal computers sometimes get in the way of "thoughtful concentration and problem solving", Abby Smith Rumsey (2016, p. 12) wrote. In electronic reading, readers develop the 'grasshopper mind', hopping from point to point (Wiegel & Gardner, 2009), getting distracted by the hypertexts and detouring from the path of linear thought, therefore affecting comprehension. "Evaluating links and navigating a path through them (hypertexts)", wrote Nicholas Carr, "involves mentally demanding problem-solving tasks that are extraneous to the act of reading itself. Deciphering hypertext substantially increases readers' cognitive load and hence weakens their ability to comprehend and retain that they're reading" (2011, p. 126).

Substantial empirical evidence exists on this subject. For instance, Canadian scholars David Miall and Teresa Dobson (2001) conducted a study in 2001 where 70 individuals were asked to read a short story. The participants were divided into two groups; one group read the story in a traditional manner, while the second group read the story with links as we would find on the Internet. Miall and Dobson found that the second group of readers took a longer time to read the story, and yet more of them reported the inability to understand the story. They were more confused about what they read as compared to the first group (for further insights into the negative impact of screen-based media on cognitive abilities, see Horowitz-Krause & Hutton, 2018; and Meri et al., 2023).

One contributing factor is that reading on the Internet is different. Research conducted by Liu Ziming (2005, 2012), a professor at San Jose State University who tracked eye movement, found that web users hardly followed a line-by-line way of reading (the "E" style). A vast majority of the respondents skimmed the text quickly and skipped lines as they went along, resembling the letters "F" or "T". The eyes would move so quickly across the screen that one could only wonder if the article was read at all or with a zig-zag style, with some reading on the first few lines of the screen, then a little in the middle and a few lines at the end. The lack of

'cognitive patience' has an adverse effect on critical analysis in the deep-reading circuit, which demand patience, time and effort.

Even avid readers are not spared. Susan Blum (2016), a cultural, linguistic, and psychological anthropologist at University of Notre Dame has said that she could not concentrate on reading as much as she used to do with all the digital devices, applications and contents. Nicholas Carr (2016, p. 231) wrote, "Immersing myself in a book or a lengthy article used to be easy. My mind would get caught up in the narrative or the turns of the argument, and I'd spend hours strolling through long stretches of prose. That's rarely the case anymore. Now my concentration starts to drift after two or three pages. I get fidgety, lose the thread, begin looking for something else to do." Bill Gates confessed in a speech his preference for printed paper to computer screens for extensive reading: "Reading off the screen is still vastly inferior to reading off paper. Even I, who have these expensive screens and fancy myself as a pioneer of this Web lifestyle, when it comes to something over four or five pages, I print it out and I like to have it to carry around with me and annotate. And it's quite a hurdle for technology to achieve to match that level of usability" (cited in Darnton, 2009, p. 69). Maryanne Wolf, author of bestsellers "Readers, come home" and "Proust and the squid" admitted that the way she read has changed over the years. "I now read on the surface and very quickly", she writes, "I read too fast to comprehend deeper levels, which forced me constantly to go back and reread the same sentence over and over with increasing frustration" (Wolf, 2018, p. 100). Imagine those who are less devoted to reading.

While it may be accurate to assert that, thanks to the Internet, people today read more than individuals did in the 1970s or 1980s, it is important to recognise that this represents a different kind of reading, underpinned by a distinct mode of thinking. It is a concern that online readers prioritise efficiency and immediacy above all else, potentially compromising our capacity for the profound reading and critical thinking that flourished in the wake of the invention of the printing press.

Deep reading, as facilitated by printed pages, holds immense value. It extends beyond the mere acquisition of knowledge from an author's words; it encompasses the intellectual resonances these words trigger within our own minds. Within the tranquil realms carved by focused, uninterrupted book reading, or any other form of contemplation for that matter, we cultivate our unique associations, formulate inferences and analogies, and nurture our original ideas. In essence, deep reading and deep thinking are inseparable companions. Notably, our ability to interpret text and form intricate mental connections, which develop when we engage in deep, distraction-free reading, is at risk of erosion.

There is another reason why print books have remained popular among the readers. The size of printed books, designed to fit comfortably in one hand, uniquely offers a distinct tactile experience compared to the elongated scrolls or rigid clay tablets of the past. Printed book lovers speak about the feel of the paper, the sound of a page turn, the broad margins to write personal comments and the thickness

of the book to remind us how much we have read and how far we are from the concluding chapter. Notably, physical books are marvels of engineering. They excel at packaging vast amounts of information within their compact pages, allowing for easy reference and navigation. Their ergonomic design makes them a joy to flip through, inviting readers to explore their content at their own pace. These books are also remarkably durable, resilient against wear and tear, and capable of withstanding the test of time without the need for upgrades or downloads. Books, wrote Martyn Lyons (2013, p. 7), “do not need batteries they do not get infected by virus and when you close a book you never need to ‘save’ because you will never lose your data”.

One of the most cherished aspects of physical books is their smell. Book enthusiasts hold this olfactory quality in high regard, often finding solace in the comforting scent of well-worn pages. This aroma is so treasured that many readers resist the allure of odourless electronic books, valuing the sensory connection that physical books provide. Bali Rai said, “No e-reader will ever replace the beauty of the fully formed, 3-D book. Technology has its place, but it would not even exist without books and libraries. I love the feel and smell of libraries” (2012, p. 123).

The assertion that reading physical books is a positive way forward has profound implications for the realm of teaching and learning. In an era marked by rapid technological advancements and the ubiquitous presence of digital media, this conclusion prompts us to reevaluate the role of physical books in education and consider the broader impacts on pedagogy, cognition, and the overall learning experience. At the heart of this assertion lies a fundamental recognition of the unique advantages offered by physical books. Unlike their digital counterparts, physical books engage multiple senses, stimulating a deeper connection with the material. The tactile sensation of turning pages, the faint aroma of ink and paper, and the weight of a book in one's hands create a multisensory experience that digital screens cannot replicate. This sensory engagement is particularly crucial for young learners who are in the process of developing their cognitive abilities.

Incorporating physical books into the learning process provides myriad benefits. As mentioned, one of the most significant advantages is enhanced comprehension and retention. The act of physically flipping pages aids memory retention, allowing students to absorb and recall information more effectively. Furthermore, the absence of digital distractions often associated with electronic devices promotes focused and undistracted reading, a key factor in comprehension.

Beyond cognitive benefits, physical books have a profound impact on the emotional and psychological aspects of learning. The tangible nature of books fosters a sense of ownership and attachment. Students develop a personal connection with their books, treating them as cherished possessions. This emotional bond can motivate students to read more extensively and with greater enthusiasm, fostering a lifelong love for literature.

Physical books also play a pivotal role in creating an optimal learning environment. For one, physical books offer a respite from the constant connectivity and screen-based activities that dominate modern life. The act of reading a physical book encourages individuals to unplug from the digital world, promoting mindfulness and focus. This separation from screens is especially vital for the well-being of young learners, who may already spend significant time engaged with digital devices. The physical presence of books, for example, in libraries, imparts a sense of academic seriousness and reverence for knowledge. Libraries offer a space for quiet reflection, concentration, and scholarly pursuits—an atmosphere that is increasingly rare in the noisy, interconnected world of digital learning.

The assertion that physical books are the way forward highlights the critical role of libraries in education. Libraries serve as repositories of knowledge, offering students access to a wide range of academic resources, including books, journals, and historical documents (Pettergree & Der Weduwen, 2021). In addition to providing access to physical books, libraries foster a culture of exploration and research. The serendipitous discoveries that can occur while browsing library shelves are invaluable to the intellectual development of students. Thus, the continued existence and support of libraries are integral to the advancement of education.

The assertion that reading physical books is the way forward also calls for a *balanced approach* to technology in education. In recent years, digital learning tools and e-books have gained prominence in classrooms. While these digital resources offer undeniable advantages, they should be integrated thoughtfully to complement rather than replace physical books. Striking this balance ensures that students benefit from the strengths of both digital and print media. Digital tools can enhance the learning experience by providing interactive and multimedia content. They offer accessibility features that cater to diverse learning needs, such as text-to-speech functionality for students with visual impairments. Digital platforms also enable real-time collaboration and communication among students and educators. These advantages make digital resources valuable components of modern education. While digital content offers convenience and accessibility, it also poses challenges related to screen time, digital distractions, and concerns about the permanence of digital formats. Physical books, by contrast, provide a timeless and enduring means of preserving knowledge. A balanced approach acknowledges that each medium has its strengths. Educators should carefully select the appropriate format for specific learning objectives. For instance, while a digital platform may facilitate collaborative group projects, a physical book may be better suited for in-depth literary analysis.

## Conclusion

This opinion piece has explored the enduring role of books as a resilient medium of instruction. Throughout history, from ancient scrolls to modern printed volumes, physical books have consistently demonstrated their importance in education and knowledge dissemination. Books have served as invaluable repositories of human thought, culture, and

progress, preserving the wisdom of the past and guiding society toward greater enlightenment. They have played pivotal roles in inspiring change, challenging norms, and encouraging generations to transcend their circumstances. Despite the digital age and the proliferation of screens and gadgets, books continue to captivate readers and learners. Their tangible presence and tactile experience offer a unique sanctuary for those seeking knowledge, serving as steadfast companions in the lifelong journey of learning. In an era of rapid technological change, physical books stand as a testament to the timeless value of written knowledge. As we embrace innovation, let us also acknowledge the timeless wisdom held within the pages of physical books, forever prepared to lead us in our pursuit of knowledge.

## References

Bickerts, S. (2006). *The Gutenberg elegies: The fate of reading in an electronic age*. Farrar, Straus & Giroux.

Blum S. (2016). *'I love learning. I hate school': An anthropology of college*. Cornell University Press.

Carr, N. (2011). *The shallows: What the Internet is doing to our brains*. W.W. Norton and Company.

Carr, N. (2016). *Utopia is creepy and other provocations*. W.W. Norton & Company.

Cohen, R. (2023). *Making history: The storytellers who shaped the past*. Weidenfeld & Nicolson.

Cornelius, J. (1991). *When I can read my title clear: literacy, slavery, and religion in the Antebellum South*. University of South Carolina Press.

Cummings, B. (2022). *Bibliophobia: The end & the beginning of the book*. Oxford University Press.

Darnton, R. (2009). *The case for books: Past, present, and the future*. Public Affairs.

Drucker, P. (1999). *Management challenges for the 21st century*. HarperCollins Publishers Inc.

Duncan, D., & Smyth, A. (2019). *Book parts*. Oxford University Press.

Duncan, D. (2022). *Index, a history of the: A bookish adventure from Medieval manuscripts to the digital age*. W.W. Norton and Company.

Faverio, M., & Perrin, A. (2022). *Three-in-ten Americans now read e-books*. Pew Research Centre. <https://www.pewresearch.org/short-reads/2022/01/06/three-in-ten-americans-now-read-e-books/>

Ferner, A. & Meyns, C. (2021). *The philosophers' library: Books that shaped the world*. Ivy Press.

Fischer, S. (2021). *A history of writing*. Reaktion Books.

Gabriel, Z. (2004). *So many books: Reading and publishing in an age of abundance*. Paul Dry Books.

GCPEA (Global Coalition to Protect Education from Attack) (2019). *"It is very painful to talk about": Impact of attacks on education on women and girls*. GCPEA.

Goody, J. (2006). *The thief of history*. Cambridge University Press.

Graeber, D., & Wengrow, D. (2022). *The dawn of everything: A new history of humanity*. Penguin Books.

Grayling, A. (2022). *For the good of the world: Why our planet's crises need global agreement now*. OneWorld Publications.

Handy, C. (1995). *The empty raincoat: Making sense of the future*. Aarow Books.

Horowitz-Kraus, T., & Hutton, J. (2018). Brain connectivity in children is increased by the time they spend reading books and decreased by the length of exposure to screen-based media. *Acta Paediatrica*, 107(4), 685–693.

King, S. (2012). *On writing: A memoir of the craft*. Hodder.

Kurzweil, R. (1999). *The age of spiritual machines: When computers exceed human intelligence*. Viking.

Liu, Z. (2005). Reading behaviour in the digital environment: changes in reading behaviour over the past ten years. *Journal of Documentation*, 61(6), 700-712.

Liu, Z. (2012). Digital reading: An overview. *Chinese Journal of Library and Information Science*, 5(1), 85-94.

Lyons, M. (2013). *Books: A living history*. Thames & Hudson.

McAllister, G., & Harati, H. (2023). Contending with controversy: Using a decision-based simulation for preservice teacher education on addressing challenged books. *Journal of Applied Learning and Teaching*, 6(Sp. Iss. 1). <https://doi.org/10.37074/jalt.2023.6.S1.7>

McLoughlin, D. (2023, January 30). *Print book sales statistics*. <https://wordrated.com/print-book-sales-statistics/>

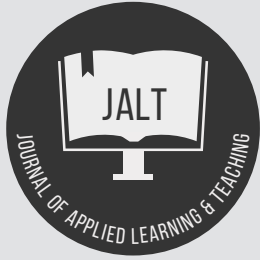
Meri, R., Hutton, J., Farah, R., DiFrancesco, M., Gozman, L. & Horowitz-Kraus, T. (2023). Higher access to screens is related to decreased functional connectivity between neural networks associated with basic attention skills and cognitive control in children. *Child Neuropsychology*, 29(4), 666-685.

Miall, D., & Dobson, T. (2001). Reading hypertext and the experience of literature. *Journal of Digital Information*, 2(1). <https://jodi-ojs-tdl.tdl.org/jodi/article/view/jodi-36>

Pettegree, A., & Der Weduwen, A. (2021). *The library: A fragile history*. Profile Books.

The Reading Agency. (2012). *The library book*. Profile Books.

- Rowe, B., & Levine, D. P. (2006). *A concise introduction to linguistics*. Pearson.
- Rumsey, A. (2016). *When we are no more: How digital memory is shaping our future*. Bloombury Press.
- Said, E. (1994). *Orientalism*. Vintage Books.
- Simonds, M. (2017). *Gutenberg's fingerprint: Paper, pixels and the lasting impressions of books*. ECW Press.
- Smith, E. (2022). *Portable magic: A history of books and their readers*. Allen Lane.
- Thomson, J. (2021). *Book wars: The digital revolution in publishing*. Polity Press.
- Vallejo, I. (2022). *Papyrus: The invention of books in the ancient world*. Hodder & Stoughton.
- Whitehead, C. (2016). *The underground railroad*. Doubleday.
- Wiegel, M., & Gardner, H. (2009). The best of both literacies. *Educational Leadership*, 66(6), 38-41.
- Wolf, M. (2018). *Reader, come home: The reading brain in a digital world*. Harper.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

## Cargo cultism and the whiteness syndrome: Fake internationalization of private universities of India

Sayan Dey<sup>A</sup>

A

*Assistant Professor, Alliance University, Bangalore*

Aritra Chakraborty<sup>B</sup>

B

*Senior Analyst, TE Connectivity, Bangalore*

### Keywords

Cargo cultism;  
internationalization;  
National Education Policy 2020 (NEP, 2020);  
privatization;  
whiteness syndrome.

### Abstract

One of the many visions of the National Education Policy 2020 (NEP, 2020), is to make efforts to internationalize the higher education system of the country. The purpose of internationalization is to make sure that like the universities from Europe, the US, Australia and other white-centric geopolitical locations of the world, the universities in India can attain global distinctions in terms of rankings, publications, curriculums and pedagogies. However, the execution process is plagued with flawed, superficial and grossly researched policies. To explain further, the initiatives for internationalizing the higher education systems are being undertaken at a rapid pace and in uncritical ways. For example, the focus of internationalization is centered on private universities, with not much focus on government-run institutions. In the name of student and faculty exchange programs, the universities in India are being flooded with white academicians whose physical visibility matters more than scholarliness. Their visibility on the university campuses is regarded as a potential marketing tool to initiate various forms of degree programs and motivate students to pay enormous amounts of registration fees. This opinion piece discusses how the project of internationalization of universities in India is engulfed with the phenomena of cargo cultism and whiteness syndrome. The arguments have been supported with personal conversations with two research participants from two private universities that are based in Noida and Hyderabad. Besides personal conversations, the arguments have also been supported by informal conversations with friends and colleagues and by analyzing photos, videos, and writings that are posted on social media by the university as markers of appreciation and success.

### Correspondence

[sayan.dey@alliance.edu.in](mailto:sayan.dey@alliance.edu.in)<sup>A</sup>

### Article Info

Received 20 November 2023

Received in revised form 7 January 2024

Accepted 16 January 2024

Available online 24 January 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.11>

## Introduction

In 2019, after submitting my PhD thesis, I (Sayan Dey) joined a private university in Noida to teach in the Department of Language and Literature. One of the many factors that motivated me to apply for the job was a set of international programs for teaching and research that looked very promising and were glamorously outlined in the university's website. I felt that such a university with multiple international collaborations will be a suitable fit for my career growth. However, after a few weeks of joining, I realized that teaching and research were regarded as the least important there. Rather, what was given more importance was, that we dressed in suits and ties; we spoke accented English (even if it sounded comical); we published our research papers in so-called international journals even if they were deemed predatory; and we had the capacity to invite white, golden-haired and blue-eyed academicians from Europe and the US. Such academicians may not have been of much relevance to the then-existing academic and research programs of the university, but they should be physically visible in the campus. Besides these, the architectural style of the university, the campus structure, the cafeterias, the classrooms and the curriculums blindly mimicked the European universities with a belief that gradually such mimicry would uplift them to the 'standards' and the 'status quo' of the prestigious Euro-North American academic institutions. With the formulation of National Educational Policy 2020 (NEP, 2020) that envisions a rigorous internationalization of the higher education system in India, these practices of cargo cultism (Feynman, 1974) have aggravated further.

During a lecture at the California Institute of Technology in 1974, Richard Feynman argued that many human communities across the planet have a fetishism towards mimicking the existential practices of certain racially privileged cultures and societies (like white-centric Euro-North American sociocultural practices) that are widely sought after across the globe. But, in reality, such fetishisms have no value except generating imaginative experiences of physical and emotional spectacularity that are underpinned with falsified promises of growth and development (Feynman, 1974). Though Richard Feynman conceptualized the notion of cargo cultism in the contexts of physics and cultural practices, the concept has been found relevant across diverse sociocultural scenarios. The phenomenon of cargo cultism is underlined by the perspective that the possession of certain objects or cargoes that were once used by the colonial officials will enhance the social, cultural, racial, and economic status of the communities, who were once colonized and enslaved by the Europeans. According to Feynman (1974), the practices of cargo cultism are usually prominent amongst the countries who have been or continue to be ridden by warfare, border disputes and refugee crises. Such situations have eventually opened up gateways for Europe and the US to physically and ideologically interfere and further enhance the already existing sociopolitical fractures. In this article, the concept of cargo cultism has been fleshed out in the context of how private universities in India mimic the Euro-North American universities physically, culturally, socially, and ideologically. The cargoes appear in the forms of Graeco-Roman architectural styles of the buildings; large theatre-styled

classrooms equipped with latest technologies; syllabuses and pedagogies that have been hijacked from various Ivy League and Russell League universities; and the functional procedures that are dictated according to the whims and fancies of certain foreign universities. The growth of cargo cultism was further intensified by the phenomenon of whiteness syndrome. Whiteness syndrome is a neocolonial social, cultural, and ideological phenomenon that seduces, motivates and convinces individuals and institutions to blindly imitate the colonially-structured, white-centric, and Euro-North American patterns of thinking and doing.

My (Sayan) experience at the university in Noida was no different. Academics from Europe, the US and Australia visited the campus for a few days; pretended to have cracked all the codes of potentialities and crises that the university had and fed the university administrators with Euro-North American-centric ideas of growth and progress, which did not fit within the local socioeconomic contexts. While walking around the campus, I (Sayan) often saw that the faculties and staff from different European universities would visit the institution; sign Memorandums of Understanding (MOUs); click elegant photos with the inhouse leadership teams for social media bytes; and hold several meetings, whose photographs were widely visible, but the content of the discussions were unknown. Prior to their departure, they also provided feedback to the leadership team, which was implemented by frantically replacing the existing syllabuses and replacing them with Eurocentric curricular and pedagogical transformations. As a part of the change, many local Indian texts on theories and philosophies, especially in the fields of Humanities (literature, cultural studies, sociology, history, philosophy, and other subject areas), were replaced by selective western texts, which were found unrelatable to the local sociocultural contexts. The process of replacement was fancifully titled as 'research and academic upgradation,' and was dumped on students and faculties, who were clueless about how to execute them. Every concern and complaint fell on deaf ears and the students and faculties were silenced by the university management with the narrative that these changes will globally enhance the international reputation of the university, without any basis. During my one year of stay in that university, I (Sayan) had closely observed how whiteness in the forms of shifts and movements of white bodies (international guest faculties and students) through the university corridors, classrooms, cafeterias, and meeting halls enhanced the marketability of the university. The physical movements of the white bodies also saw the simultaneous movement of white-centric ideologies that make the collaborations and exchange programs available only to a fixed set of faculties and research disciplines that are usually considered to be financially competent and can generate satisfactory revenues. So, even though the MOUs were advertised as a pan-University vision, in reality, with the hidden consent of the foreign partner universities, they were made available exclusively to the faculties of engineering, technology, and business studies.

The faculties from the humanities and social sciences were completely left out, as they were found not intellectually rich enough for academic and research growth. Along with these aspects, the epidermalization of knowledge-making spaces (Fanon, 1967; Browne, 2009; Dey, 2022) in India, where the

presence and the narratives of the white and white-centric ideologies and intellectualities are considered by default relevant and superior to the knowledge systems of one's own, need critical attention. This article, through various evidences and analyses of the perspective of the private universities in India, underlines "how whiteness underwrites systems of racial oppression and how it is reproduced" within the contemporary systems of knowledge production (Owen, 2007, p. 203). The whiteness syndrome of various private universities unfolds the ways in which whiteness functions as a "structuring property of racialized social systems" (Owen, 2007, p. 203). During my stay in that institute, I (Sayan) also observed that only white academics were invited and when I (Sayan) raised this aspect with one of my colleagues, he winked at me and shared "whiteness sells" (Participant A, Personal Conversation, 2019, Noida). These experiences were not only personal but collective as well, which I realized through personal conversations with colleagues based there and in other private universities across India.

Prior to progressing with the arguments, we would like to clarify here why this article is specifically focused on private universities. We chose private universities because of our physical accessibility and work experiences. Since our first jobs, we have been working in various private institutions in India. As a result, we are highly aware of the various intentions and ideologies through which the institutions function. Also, with respect to the vision of National Education Policy 2020 (NEP, 2020), the process of internationalization of the academic system has been strategically centered on the private universities and government-run educational institutions are deliberately and systemically ignored. It is so because, in this way the higher education system of the country can be centrally corporatized and the students can be helplessly exploited in terms of admission fees, registration fees for study abroad programs, international summer schools, canteen fees, and various other ways, without any restrictions and regulations. The process of exposing the learners to various international academic and research programs is highly appreciable. However, the focus of introducing such programs in a lot of corporatized universities in India is less on scholarship building and more on profitmaking, through blindly implanting the functional procedures of the universities based in Europe and the US. This systemic ignorance of certain educational institutions at the cost of others is one of the many underlying agendas of internationalization.

With respect to these arguments, the article is divided into five sections. The introduction sets the pace of the arguments by discussing the backdrop from which the thematic and theoretical arguments have emerged in this article. The introduction is followed by the literature review that reflects on the different works that have been generated so far with regard to the overarching theme of this article. This section also highlights the scantiness of available documents. The third section reflects on the research methods of walking interviews and a close analysis of the social media pages of the universities in which the research participants are based, and how the responses have shaped the arguments. With respect to the conversations, the fourth section analyses and discusses how the conversations unfold the normalization of cargo cultism and whiteness syndrome as shortcut ways

of catching up with the West. The final section summarizes the article and recommends possible ways in which effective policies can be framed and implemented in order to internationalize the higher education system of India with the most diverse and inclusive patterns.

## Literature review

Apart from newspaper articles and various op-eds, no documents exist that particularly focus on the problematic ways in which the higher education system is being internationalized and privatized. In fact, a lot of these op-eds like "Evolving Concept of Internationalization in Indian Education" (2021) by Anirban Chakraborty, "Internationalization of Higher Education in India" (2023) by Iqbaljeet Singh Bains, and many others not only appreciate the National Education Policy 2020 (NEP, 2020) of internationalization, but also focus their arguments on how internationalization should be led by STEM (Science, Technology, Engineering and Mathematics) education. These articles also deliberately do not focus on the ways in which the disciplines of humanities and social sciences are being systemically left out of the projects of internationalization. These articles also appreciate the rapid transformations that are being brought by the universities through recruiting foreign full-time and visiting faculties. However, deliberately or ignorantly, what these articles fail to highlight is how often the quality of the foreign faculties is way below the expected standards of teaching and learning that have been discussed in the consequent sections. To elaborate further, as discussed in the fourth section of the article, the international faculties often fail to fulfil the basic requirements of an institution like framing innovative curriculums, publishing in prestigious journals with high impact factors, carrying out basic administration responsibilities as outlined in the job contracts, and others. Despite their failures, their job remains secured and often when the foreign faculties decide to leave, they are influenced to stay back by increasing their pay scale for nothing (Participant B, Personal Conversation, 2021, Hyderabad).

Amidst censorship and surveillance by the present right-wing government of India, there are very few articles like "Interrogating the Internationalization of Indian Higher Education" (2023), which openly critiques the higher education system of India. This commentary by Ranjan & Hameed focuses on "flawed assumptions" and "inherent contradictions" (2023, p. 28) of the National Education Policy (NEP, 2020) with regard to the state of the internationalization of the higher education system in India. The article does not specifically focus on the aspects of cargo cultism and whiteness syndrome, but the arguments about questionable teaching, learning and ranking policies (argued in the following sections) as shaped by the political and commercial stakeholders' critiques of the fake projects of internationalizing the Indian academia. Apart from these documents, it is also crucial to talk about National Education Policy 2020 (NEP, 2020). With regard to the aspect of internationalization, the educational policy promises to make the "Indian education system self-reliant and compliant to global standards & norms which would enable



India to attract a greater number of students from abroad" (National Education Policy, 2020). The policy also outlines the methodologies that would be used to implement the proposal like research and teaching collaborations with and "faculty/student exchanges with high-quality foreign institutions, and mutually beneficial MOUs with foreign countries" (National Education Policy, 2020). The UGC Chairman also publishes short writeups and videos on social media channels for the students, parents and faculties to clarify how the policies of internationalization would be implemented. However, in reality, the verbal assurances and blueprints hardly match with the systems in which they are implemented. As mentioned in the introductory and the following sections, these policies are nothing more than eye-washing tactics of blanketing the real intention of privatizing the education system, making it blindly and seductively white-centric, and limiting academic accessibility only to the socioeconomically privileged communities.

## Research methods

The arguments in this article have been shaped through personal conversations and close analysis of photos, videos, and descriptive texts as found in the social media pages of the universities in which the participants are based. The conversations, photos, videos, and descriptive texts have been analyzed through Critical Discourse Analysis (CDA). CDA argues that words are never neutrally produced. They convey "how we see ourselves, our identity, knowledge, values and beliefs. They are politicized even if we are not aware of it because they reflect the interests of those who speak" (Cervera et al., 2006, p. 10). CDA has been used by investigating how social dynamics and power structures are constructed within the higher educational institutions of India through implementing glamorous and enticing marketing policies across their respective campuses and social media channels. The conversations for this project took place through the research method of walking interviews, which acknowledges walking as an "experience, source of knowledge, personal sharing, and memory" (O'Neill & Roberts, 2020, p. 1).

## Participants and ethics

The conversations took place with two participants – one each from Noida and Hyderabad. The two participants were chosen on the basis of personal connections, consent and availability. Due to ethical concerns, the original names of the participants have not been revealed. The participant from Noida has been named Participant A, and the participant from Hyderabad has been named Participant B. Noida is a city, which is located in the Gautam Buddha Nagar district in the northern Indian state of Uttar Pradesh, and Hyderabad is the capital city in the southern Indian state of Telangana. It is also for ethical reasons that the photos, videos and descriptive texts have not been referenced or very specifically discussed in this article.

## Processes in which the methods have been implemented

All the conversations took place in person and inside the respective university campuses of the participants. While conversing, we walked through the open spaces, cafeterias, and corridors of different subject departments and talked about the social dynamics and power structures of different academic disciplines. We also talked about how certain academic disciplines gain validity for international accreditation and recognition (Fleming et al., 2021; Andrew, 2023). During the conversations, the physical movements across the different locations of the campus interwove with the narratives of intellectual violence, fetishism towards white-skinned academics, hierarchies in pay structures and the exclusionary functionalities of different academic disciplines. These interventions unveil diverse patterns of social practices and social relationships (Fairclough et al., 1997), and the ways in which the colonial dynamics of racial, cultural, social, and geopolitical superiorities are rekindled through localized forms of socially, culturally, and economically exclusionary systems of knowledge dissemination in contemporary India. The time period of the research was January 2019 to August 2023. The conversations that are centered on the two research participants took place between 2019 and 2021. The rest of the time period has been utilized through informal conversations with friends and colleagues and closely analyzing the social media pages of different private universities.

Besides conversations, the arguments in this article have also been framed through analyzing the social media pages of the respective universities of the research participants on LinkedIn and the patterns in which they promote their so-called innovative internationalization strategies. During the analysis, we took into consideration several photos, videos and descriptive texts that have been posted for promotions and how they exclusively focus on the appreciation by foreign academicians. We also noted that in their reflections, the foreign faculties, rather than talking about scholarship developments and capacity building of students and staff, mostly talk about the architectural aesthetics of the university campuses and the way in which they resemble the universities across Europe and the US. Such reflections further enabled us to identify the normative and systemic ways in which the performances of cargo cultism and whiteness syndrome are systemically maintained.

## Analysis and discussions

The first conversation took place in 2019 with Participant A in a private university located in Noida that boasts several hundred MOUs, faculty and student exchange programs, and varied national and international awards. Participant A was a 37-year-old male, who was then based in the Department of Communication. While conversing about the national and international prospects of his university, he shared: "Our university is being internationalized in a very hierarchical and abusive manner. On the one side, the pay scale and other facilities of the local Indian faculties are stagnated, and on the other hand, huge money is spent to bring underqualified international faculties as full-timers and guests" (Participant A, Personal Conversation, 2019, Noida). He also added: "It

is only a selective group of faculties from a set of selective disciplines, who owns the privilege of interacting with the foreign faculties. The rest of us only watch them from a distance" (Participant A, 2019). This conversation took place immediately after I joined the university. After a few months, my personal experiences were quite relatable. Participant A also mentioned that "if you see white academicians inside the campus, they may not necessarily be faculties or staff from foreign institutions. They can also be random tourists, who are allowed to freely roam around, so that the people from the university media team can film their photographs, curate false stories of international visiting staff, and proudly flaunt them across social media. Such superficial approaches towards internationalization are institutionally recognized through ranking systems and affiliations by the University Grants Commission (UGC). UGC is the highest body, which is responsible for coordinating, determining and maintaining standards of higher education. The celebration of colonial hangovers (Lele, 2012; Dey & Alamman, 2021) are motivated by the cargo-cult perspective that the random mimicry of white bodies and ideologies would enhance the intellectual impression of the university globally, which did not happen to date.

Our informal conversations with various faculties and staff from the same university revealed how these problems keep on persisting due to a lack of collective resistance, due to fear of penalization and expulsion on the one side and the seduction for being the 'wannabe European intellectual' on the other. During the conversations, many faculties shared that even though they were frustrated for not being considered intellectually capable of interacting and participating in research exchanges with foreign faculties, they did not resist because they did not want to let go of the colonially influenced desire of becoming like a white intellectual. When one looks at the videos, audios and descriptive texts on social media that promote the international programs of the university, one is bound to have a completely contradictory and misleadingly positive impression about the university.

Almost a similar experience was shared by Participant B, a 34-year-old man, who was working as an Assistant Professor in the School of Liberal Arts. With respect to the strategies of internationalization, he shared: "Almost every month we are informed about the recruitment of adjunct faculties and distinguished professors (who are by default white-skinned) from different universities of Europe and the US, and some of those universities and departments are not even prominently recognized in their own countries" (Participant B, 2021, Personal Conversation, Hyderabad). He also shared that "many professors who have been recruited did not turn up to date and many have requested to withdraw their names. And many professors who turned up did not appear to add a lot of value to the existing academic and research structure of the institution" (Participant B, 2021). Moreover, as the participant shared, the university continues to claim a lot of white foreign faculties as their own, even after they have officially resigned. He also added: "When international faculties arrive on the campus, their itineraries are set up in such a manner so that they do not have much scope for public interactions and remain only limited to a group of faculties and staff. They always receive 'special

treatment' from the management because of their superior class status and foreign academic degrees" (Participant B, 2021). Besides him, the informal conversations with other faculties revealed that the practices of cargo cultism and the whiteness syndrome in terms of Euro-North American-centric architectures, body languages, behavioral patterns, and dressing styles are widely encouraged. These white Euro-North American-centric aspirations are strategically and institutionally preserved as smartness, professionalism, and marketing gimmicks. These fetishisms toward whiteness and foreignness are other instances of how the colonial hangover in the forms of celebrating white physical and ideological presence in the university campuses serve as one of the many foundational dimensions of internationalizing the higher educational institutions in India. At the time of sharing their experiences, both the participants also revealed how their respective institutions blindly mimic the architectural styles of European and American universities to make the students and staff feel as if they are studying or working in a Euro-North American University.

These initiatives convert universities from sites of intellectual and critical thinking towards an imaginative promised land of happiness (Mignolo, 2007), where the knowledge-making processes are entirely focused on advertising, marketing, and the salability of knowledge, power and being (Ndlovu-Gatsheni, 2015; Mbembe, 2016) that are manufactured and globally deployed as authentic and impactful by Europe and the US. Having read the arguments and the analyses in this article, the readers might feel skeptical about how we derived our conclusions just based on two personal conversations. Therefore, we would like to clarify here that along with the two conversations, the arguments have also been shaped through personal experiences and the informal interactions that we had with various friends and colleagues, who are based in different private universities across the country. Also, as mentioned in the research methods section, the analysis of photos, videos, and descriptive texts has also enabled us to unfold the phenomena of cargo cultism and whiteness syndrome through diverse contexts and situations.

## Conclusion and possibilities

Altogether, this article makes an effort to unpack the ways in which genuine scholarly, critical, and intellectual developments are being habitually compromised through rapidly reducing the knowledge-making processes of teaching, learning and researching into a mere state of gimmickry. Within a state of gimmickry, serious thinking and discussions are erased and replaced by performances of spectacularity that are spearheaded by white, heteronormative and Euro-North American-centric bodies, ideologies and knowledge structures (Dey, 2020). The uniqueness of this article lies in the aspect that it exposes the underlying fakeness and capitalistic propaganda of internationalization of many private universities in India. No research article has been written so far that critically and elaborately reflects on how many higher educational institutions conceptualize the project of internationalization as a fraudulent money-lending scheme that lures students with promises of 100 per cent campus placements, attractive pay packages and

professional infrastructural support. Such promises often turn out to be false and exaggerated in nature and by the time the students finish the courses, they feel lost, deceived and exhausted. As discussed in this article, the pan-Indian project of internationalizing the higher education system, rather than engaging with genuine research works, critical discussions, and innovative curriculums and pedagogies, are more focused on creating fake templates of global education and knowledge values in the forms of spectacular photos, videos, architectures and resource systems that are visually appealing, but ideologically hollow.

It is important to note that the criticisms of fake internationalization, intellectual cargo cultism, and whiteness syndrome in this article are not intended to dismiss the values of internationalization as a whole. The real intention is to urgently expose the hidden intentions of education ministries and organizations in India to restrict access to education, learning and intellectual growth within the socioeconomically privileged communities. It is so, because, such communities are always ready to celebrate the privatization and commercialization of the education system without questioning and critiquing the process. The phenomena of cargo cultism and whiteness syndrome also remind us how, despite several years of India's judicial independence from the British, the specters of colonization continue to haunt the country by stealthily invading the individual psyches, societies, and the knowledge-producing systems, and successfully convincing the people to celebrate the flawed, white-centric, and invasive ideological and epistemological approaches of many Euro-North American educational institutions. However, we cannot conclude our arguments through criticisms and lamentations. Along with criticisms, it is crucial to find possible pathways for rectifying the flawed internationalizing approach of the private higher educational institutions in India. One of the possible initiating points could be to thoroughly revise the national educational policies and clearly state that internationalization should not be limited to the architectural innovations of university campuses, flooding the campuses with white academicians, and signing MOUs. Educational institutions need to engage in genuine scholarship growth through curricular, pedagogical, and research collaborations by inviting young and experienced researchers and professors (irrespective of their caste, class and race) from reputed international universities with cutting-edge research and teaching-learning experiences, and encouraging in-house students and staff to collaborate with them in unbiased ways.

Another possible way of countering fake internationalization could be to allot a significant part of the education budget towards equal capacity building in all the academic and research disciplines. Usually, the educational ministry centrally encourages students and staff to engage in research projects that are centered on the fields of business, science, technology, engineering and management. A small portion of the budget is allotted to humanities and social sciences, and only for certain specific subjects like political sciences, international relations, and economics. The educational ministry of India mostly signs MOUs with educational ministries from Europe, Australia and the US. Also, many times, the MOUs are signed to enhance the capacity building of students and staff who are only associated with

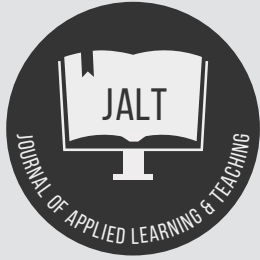
STEM education (Tilak, 2021). This problematic bias needs to be dismantled and research and teaching in the fields of humanities and social sciences need to be encouraged by allotting sufficient grants for teaching, research, and publication developments through constructing research centers and incubation hubs; building open-access research publication models; incentivizing students and staff for producing pathbreaking researches, designing residential fellowship programs to encourage international researchers and thinkers, and in various other ways. Along with these initiatives, it is essential to build educational and research bodies with experienced academicians and researchers, who can keep track of the predatory internationalizing practices of the educational institutions in the forms of recruiting foreign faculties, publishing in international journals and signing exchange programs, and alert the public by exposing them. Already organizations like India Research Watchdog (access link: <https://www.linkedin.com/company/india-research-watchdog/>) and Retraction Watch (access link: <https://www.linkedin.com/company/retractionwatch/>) have been taking initiatives to investigate and expose the fake international agendas of different private universities in India by revealing the names of the individuals and organizations and justifying their predatory acts of recruitments, publications, and teaching-learning models through thoroughly researched data and statistics. Though 'Retraction Watch' is a US-based organization, they have designated research and investigation teams in India, who keep a thorough track of predatory practices of internationalization across public and private universities in India. These initiatives are effective, but they may not be sufficient because with just two to three organizations, it is difficult to take into account the academic and research activities of 455 private universities in India. Therefore, more such organizations are required.

In addition to these institutional initiatives, it is crucial for teachers, learners and researchers to build self-consciousness. The presence of a few photos and videos of white foreign academicians on social media pages and institutional websites should not be regarded as parameters of high intellectual scholarliness. As discussed in the previous sections of this article, such representations can be extremely deceptive and misleading. So, prior to taking admissions or seeking jobs, individuals should thoroughly research their chosen institution by getting in touch with faculties and students, looking for performance statistics of the students, and knowing about the research and teaching portfolios of the faculties. However, these possibilities and recommendations are not the ultimate ones and this article serves as a warm invitation to scholars across India to continue with these discussions and share other potential recommendations. As authors, we firmly believe that this article will function as an efficient resource to encourage researchers to produce more such works that would expose the fake strategies of growth and development of the higher educational institutions in India and generate awareness amongst the learners and the teachers.

## References

- Andrew, M. B. (2023). Come to the cabaret: Voices from the modern university. *Journal of Applied Learning & Teaching*, 6(2), 17-27. <https://doi.org/10.37074/jalt.2023.6.2.19>
- Bains, I. S. (2023). *Internationalisation of higher education in India*. Invest India. <https://www.investindia.gov.in/team-india-blogs/internationalisation-higher-education-india>
- Browne, S. (2010). Digital epidermalization: Race, identity and biometrics. *Critical Sociology*, 36(1), 131-150. <https://doi.org/10.1177/0896920509347144>
- Cervera, J. T., Postigo, M. L., & Herrero, R. D. (2006). What is critical discourse analysis? *Quaderns de Filologia*, 11, 9-34.
- Chakraborty, A. (2021). Evolving concept of internationalisation in Indian higher education. *Times of India*. <https://timesofindia.indiatimes.com/readersblog/anirbanspeaks/evolving-concept-of-internationalisation-in-indian-higher-education-39278/>
- Dey, S., & Alamman, P. (2021). 'Covid batch': A case study on unethical assessment practices in selected higher educational institutions in Assam and West Bengal, India. *Journal of Applied Learning & Teaching*, 4(2), 130-134. <https://doi.org/10.37074/jalt.2021.4.2.11>
- Dey, S. (2020). In search of the de-colonial turn in the Indian academia: De-colonising the philosophies of knowledge production – a mult-versal shift. In B. Mpofu & S. J. Ndlovu-Gatsheni (Eds.), *The dynamics of changing higher education in the global south* (pp. 54-75). Cambridge Scholars Publishing.
- Dey, S. (2022). *Green academia: Towards eco-friendly education systems*. Routledge India.
- Fairclough, N., Mulderrig, J., & Wodak, R. (1997). Critical discourse analysis. In T. A. Van Dijk (Ed.), *Discourse studies: A multidisciplinary introduction* (pp. 258-274). Sage. <https://doi.org/10.4135/9781446289068>
- Fanon, F. (1967). *Black skin, white masks*. Grove Press.
- Feynman, R. P. (1974). *Cargo cult science*. Caltech's Library. <https://calteches.library.caltech.edu/51/2/CargoCult.htm>
- Fleming, P., Rudolph, J., & Tan, S. (2021). 'Never let a good crisis go to waste'. An interview with Professor Peter Fleming on dark academia, the pandemic and neoliberalism. *Journal of Applied Learning & Teaching*, 4(2), 110-120. <https://doi.org/10.37074/jalt.2021.4.2.14>
- Lele, S. (2012, June 20). *Colonial hangover persists*. Down To Earth. <https://www.downtoearth.org.in/blog/colonial-hangover-persists-38450>.
- Mbembe. A. (2016). Decolonizing the university: New directions. *Arts and Humanities in Higher Education*, 15(1), 29-45. <https://doi.org/10.1177/1474022215618513>
- Mignolo, W. D. (2007). Delinking: The rhetoric of modernity, the logic of modernity and the grammar of de-coloniality. *Cultural Studies*, 21(2), 449-514. <https://doi.org/10.1080/09502380601162647>
- Ndlovu-Gatsheni, S.J. (2015). Decoloniality as the future of Africa. *History Compass*, 13(10), 485-496. <https://doi.org/10.1111/hic3.12264>
- NEP National Education Policy 2020. (2020). *Internationalization of higher education*. Ministry of Education. <https://www.education.gov.in/nep/internationalization-education>.
- O'Neill, M., & Roberts, B. (2020). *Walking methods: Research on the move*. Routledge.
- Owen, D. S. (2007). Towards a critical theory of whiteness. *Philosophy & Social Criticism*, 33(2), 203-222. <https://doi.org/10.1177/0191453707074139>
- Ranjan, R., & Hameed, A. (2023). Interrogating the internationalisation of Indian higher education. *Economic & Political Weekly*, 58(17), 28-32.
- Tilak, B. G. J. (2021). Students' perspectives on quality of engineering education in India. *Journal of Applied Learning & Teaching*, 4(1), 56-71. <https://doi.org/10.37074/jalt.2021.4.1.15>

Copyright: © 2024. Sayan Dey and Aritra Chakraborty. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

---

## Rise of the robots: What it means for educators

---

Augustine Osamor  
Ifelebuegu<sup>A</sup>

A

*Professor, Institute of Education and Life-long Learning, Victoria University, Kampala, Uganda*

---

### Keywords

Artificial intelligence;  
chatbots;  
educators;  
future education;  
generative AI;  
robot teachers.

---

### Correspondence

[augustine.ifelebuegu@vu.ac.ug](mailto:augustine.ifelebuegu@vu.ac.ug)<sup>A</sup>

---

### Article Info

Received 22 October 2023  
Received in revised form 30 January 2024  
Accepted 13 February 2024  
Available online 16 February 2024

**DOI:** <https://doi.org/10.37074/jalt.2024.7.1.16>

---

### Abstract

The proliferation of artificial intelligence (AI) and robots is precipitating profound shifts in several sectors, including education. A pressing inquiry arises among scholarly communities: To what extent will the advancement of technology supplant or surpass the conventional responsibilities of educators? The process of dissecting and analysing this subject is intricate, with many layers to consider. Demonstrating this phenomenon, a British educational institution has recently designated an artificial intelligence (AI) robot named Abigail Bailey as one of its "co-headteachers." Abigail possesses sophisticated machine learning skills that enable her to process extensive information efficiently, akin to AI systems such as ChatGPT that simplify user interactions through algorithmic-generated answers. This breakthrough raises the question: Does this mark the beginning of a period in which robots assume a prominent position in education, hence threatening the traditional duties of educators? This opinion piece explores the essential factors that need to be taken into account.

## Introduction

The landscape of education is witnessing a revolutionary change with the advent of robotic teachers and bots, reshaping the traditional methods of teaching and learning. These technological marvels, born out of the marriage between artificial intelligence (AI) and robotics, are not just futuristic concepts but practical tools increasingly being incorporated into classrooms worldwide. Understanding what they are and the role they play is crucial in comprehending their future trajectory in education.

Ro(bots) are physical machines that can interact with the physical world. They often have mechanical components like arms, wheels, or other forms of mobility and can perform a variety of tasks ranging from industrial manufacturing to household chores (Catlin et al., 2018). Bots, short for robots, in the context of computing, are software applications that perform automated tasks over the internet. Unlike physical robots, they exist purely in the digital realm.

Modern robotic teachers are physical entities, often resembling humans, equipped with AI technology. They are designed to interact with students, deliver educational content, and perform tasks traditionally associated with human teachers (Tang et al., 2023). Unlike static robots of the past, these advanced machines can move, gesture, and exhibit a range of facial expressions, making them more relatable and engaging for students. Chatbots (a type of bot), on the other hand, are AI-driven software programs designed to simulate conversation with human users. In educational settings, they function primarily on computers or mobile devices, providing students with a conversational interface for learning. Chatbots can answer questions, provide explanations, and guide students through learning materials, much like a personal tutor.

The AI programs powering the robots and chatbots use natural language processing (NLP) to understand and respond to student queries. They can handle a range of tasks, from assisting with homework and grading assignments to facilitating revision and offering additional learning resources. Their ability to provide immediate, on-demand assistance makes them an invaluable resource for students. The AI programs enable robots to learn and adapt to the needs of individual students. They can assess students' strengths and weaknesses, customize teaching methods, and provide personalized feedback (Ifelebuegu et al., 2023). For instance, a robotic teacher in a language class can converse with students, correct their pronunciation in real-time, and adapt conversations to suit each learner's proficiency level.

This marriage of AI and robots is playing a central role in driving this transition, presenting potential opportunities for a customized and personalized learning experience (Edwards & Cheok, 2018; Huang et al., 2023; Ifelebuegu et al., 2023). The utilization of robots as tutors represents a significant advancement in the field of educational technology. According to Rosanda and Istenic Starcic (2019), advanced AI algorithms enable the provision of tailored educational experiences by robotic tutors. Educators have the ability to assess a student's ongoing development, identifying their

aptitudes, as well as areas in need of improvement, in a timely manner. This enables them to customize instructional approaches and content to suit the individual student's needs. The provision of personalized attention at this level guarantees that every student is able to advance at their most effective rate, potentially narrowing the disparity between high-achieving students and others who may require further assistance. However, similar to other technical advancements, the integration of AI robots in education presents a range of problems and potential benefits (Grace et al., 2018; Green et al., 2019; Hu et al., 2023).

The applications of AI robots within the field of education are not only a hypothetical notion, but rather an emerging actuality. A notable recent advancement in this field occurred in October 2023 with the incorporation of artificial intelligence (AI) into the administrative framework of Cottesmore School, a United Kingdom educational institution in West Sussex. This integration has resulted in the appointment of an AI robot named Abigail Bailey as a 'co-headteacher'. The robot has been bestowed with a name as a tribute to a former headmistress, and it provides assistance to the present headmaster across several domains. Abigail Bailey possesses superior machine learning abilities, which facilitate the efficient analysis of large-scale datasets. This capability mirrors the operations of chatbot services, like ChatGPT, where users interact and receive feedback from the system. Previously, a humanoid robotic teacher called "Musio", which uses AI to engage in naturalistic conversations in English, has been used to improve students' language skills (Hooper, 2018). Similarly, "Pepper", a creation by SoftBank Robotics, has been employed as a teaching assistant, demonstrating the capability to engage students through interactive learning activities (Pandey & Gelin, 2018).

Within contemporary academic discourse, a pertinent inquiry reverberates: In light of rapid technological advancements, is it conceivable that robots may ultimately replace human educators? The response to the question is intricate and encompasses several aspects (Virgillito, 2017; Selwyn, 2019). As these innovations make inroads into education, there is a growing curiosity about whether future classrooms might lack human touch. In historical context, with the advent of the printing press, scribes feared unemployment. However, the opposite transpired: literacy soared, and the number of writers multiplied. Should today's educators harbor similar apprehensions about their professional relevance? This article explores the changing interactions between robots and human teachers in the future educational scene.

## AI robots in education

The subjects of "Human versus Robot," "Machine versus Man," and "Traditional Pedagogy versus Digital Automation" have been the focus of scholarly discussion in recent years (Virgillito, 2017; Sperlinger, 2020; Johnson & Acemoglu, 2023; Joseph et al., 2024). The concept of robots replacing human positions has elicited a range of perspectives, both in cinematic representations and scholarly and professional discussions. Considering the ongoing technological advancements, namely in the fields of AI and robotics, it is

improbable to foresee education being unaffected (Mubin et al., 2013; So & Lee, 2023). Goldman Sachs predicts the potential loss or diminishing of over 300 million jobs if AI lives up to its current hype (Kelly, 2023).

Undeniably, robots present a series of advantages in the educational landscape. They have the capability to offer bespoke learning pathways, meticulously analysing a student's proficiencies and areas of development to modulate lessons instantaneously. This custom instruction can significantly enhance learning outcomes, particularly for learners who might not excel within the conventional educational frameworks (Chen et al., 2023; Ifelebuegu, 2023). Furthermore, in geographic locations that grapple with a dearth of educators or subpar educational resources, robotic solutions could bridge these vast divides, ensuring uniform and quality instruction.

A multitude of research highlights the advantageous impacts of AI-powered robotic instructors on student academic achievement. According to existing research, robotic teachers have the potential to enhance student connection, motivation, and engagement within a learning setting. By providing students with a personalized and adaptable learning experience, these robots enable them to progress according to their own requirements and preferences. In addition, the fast feedback provided by AI robotic educators enables students to promptly identify and correct faults. The use of AI robotic educators can enhance a student's academic experience by incorporating captivating and immersive educational approaches, resulting in a more efficient and effective learning environment (Robertson, 2022; Ifelebuegu et al., 2023).

With the promise of delivering personalized learning experiences and the potential to break down geographical and economic barriers in education, the implications of AI robots in teaching and learning are profound. As we contemplate the future of these technological marvels in the classroom, it is essential to project both their transformative potential and the challenges they might introduce.

### **Robotic tutors and personalized learning**

One of the most groundbreaking applications of robots in education is their role as tutors. Using sophisticated AI algorithms, these robotic tutors can deliver personalized learning experiences. They can analyze a student's progress, strengths, and areas of struggle in real time, tailoring lessons accordingly (Conti et al., 2020). The provision of personalized attention at this level guarantees that every student is able to advance at their most effective rate, potentially narrowing the disparity between advanced learners and those who may require further assistance.

### **Robots as learning companions**

Beyond the traditional tutor-student dynamics, robots are emerging as learning companions (Aziz & Ghanimi, 2020; Cagiltay et al., 2022, Huang et al., 2023). These interactive and engaging robots have the potential to foster a

student's curiosity and enthusiasm for a certain subject. The use of gamification through robotics can serve as a powerful tool for students by transforming the learning process into an engaging and participatory experience. In their latest publication, Huang et al. (2023) introduced an innovative framework known as iSTAR (Intelligent Human-Machine Synergy in Collaborative Teaching), which focuses on the integration of human and machine collaboration in educational settings. This framework underscores the combined use of three distinct technologies: digital twins, avatars/agents, and physical robots, to enhance the teaching and learning experience. These technologies can work collaboratively with human educators, offering a multi-faceted approach to educational delivery.

### **Special needs and therapeutic education**

Robots have demonstrated significant potential in facilitating support for special needs. Robots have been found to offer a valuable means of facilitating social skill development and repeated learning activities for students diagnosed with autism (Alabdulkareem et al., 2022). Several researchers have explored the applications of socially assistive robots, including humanoids, in teaching and learning (Papadopoulos et al., 2020). By providing constant and non-threatening contact, robots can effectively support students in their developmental journey. Due to their consistent behavior, they serve as an optimal learning companion for students who may encounter difficulties or feel overwhelmed by human interactions.

### **Practical skill development**

The 21st century has also seen robots playing a role in cultivating practical skills, especially in the realm of STEM (Science, Technology, Engineering, and Mathematics) education. Robots can offer hands-on learning experiences, allowing students to understand complex concepts by seeing them in action (Greca Dufranc et al., 2020). Robotics kits and platforms enable students to construct, program, and experiment, hence promoting critical thinking, problem-solving, and engineering skills. According to Nourbakhsh (2015), the utilization of educational robots as teaching aids has significantly influenced innovation and the development of learners' talents, particularly in STEM subjects.

### **Overcoming geographical and economic barriers**

One of the most transformative potentials of robots in education is their capacity to democratize learning. In regions where access to quality education or qualified teachers is limited, robots can step in, offering consistent, high-quality instruction. In addition, the availability of online connectivity enables a robot situated in a remote village to possibly have access to high-quality instructional materials, levelling the playing field for students worldwide (Edwards & Cheok, 2018).

Although robots have great promise in the field of education, it is important to note that teaching encompasses more than just the transfer of knowledge. It involves the provision of mentorship, emotional assistance, and the development of interpersonal abilities. Human educators have a crucial and influential role in the emotional and psychological growth of students. They serve as a source of inspiration and motivation and possess a deep understanding of the complex intricacies of human behavior and emotions. The existing capabilities of robots do not extend to encompass emotional intelligence, which involves the capacity to empathize, motivate, and establish human connections. The next section expands further on the key values that human educators bring to the education scene.

### **The unique value of human educators**

Human educators, with their unique blend of skills and qualities, offer something to the classroom that AI robots cannot yet emulate. Their capacity to empathize with students, adapt teaching methods based on individual needs, offer personalized guidance, and nurture critical thinking are indispensable aspects of holistic education. These attributes, which revolve around emotional intelligence, creativity, and adaptability, set human educators apart from their AI counterparts (Selwyn, 2019).

Undoubtedly, AI robots exhibit considerable promise in enhancing the educational domain, yet it is important to acknowledge that they are not devoid of constraints. Understanding intricate human emotions or engaging in profound social exchanges can be challenging for AI. The profession of teaching encompasses more than the mere dissemination of knowledge; educators frequently assume the roles of mentors, advisers, and emotional support for their students. The assertion that AI robots might potentially supplant human educators oversimplifies the complex and diverse responsibilities that instructors have (Han et al., 2023; Tlili et al., 2023, Rudolph et al., 2023).

In addition to the transmission of knowledge, instructors assume crucial duties in fostering inspiration and providing mentorship to students. Robots are unable to achieve the authentic human connection, characterized by their intricate emotional subtleties and capacity to evoke intense emotions. While certain administrative and repetitive tasks within education can be delegated to automation, this does not lessen the importance of teachers. Instead, it refines their role, potentially giving educators more opportunities for research, creative pedagogical innovations, and fostering deeper student relationships.

At the core of education, teachers emerge not just as knowledge facilitators but as guiding beacons, sometimes even echoing parental sentiments within the educational realm. As education undergoes metamorphoses with advancing times, recognizing and honoring the diverse, influential roles educators undertake becomes even more crucial. They do not merely teach; they shape futures. And in this intricate endeavor, while robots can be invaluable aides, they cannot replace the human essence of teaching. In light of societal evolution and the dynamic nature of educational

approaches, it is crucial to comprehend the intricate and diverse responsibilities that instructors assume in shaping the future of their students and, therefore, the global landscape. Despite the rapid innovation and proliferation of AI in educational tools, the World Economic Forum forecasts a 10% increase in jobs within the education sector by 2027 (see Rudolph et al., 2023). This prediction emphasizes the enduring and essential role that human teachers play in the field of education.

### **The torchbearers of knowledge**

At the most fundamental level, educators bear the responsibility of transmitting knowledge. This encompasses more than just memorization or traditional instructional methods; it involves promoting comprehension of intricate ideas, stimulating inquisitiveness, and cultivating a sincere passion for the discipline. According to Hattie (2009), this essential role transcends the mere dissemination of facts; it involves molding character, fostering curiosity, and igniting a lifelong passion for learning. Teachers, as described by Hattie in his seminal work "Visible Learning," are at the heart of the educational process, playing a pivotal role in influencing student outcomes. A mathematics educator, for example, not only imparts numerical knowledge but also fosters the development of analytical reasoning skills. A literary instructor not only engages with canonical works but also fosters an awareness of artistic expression, cultural understanding, and the complexities of human emotions.

Standing at the helm of the educational journey, teachers act as custodians and conveyors of knowledge, bridging the gap between the wealth of past wisdom and the challenges of a rapidly changing future (Freire, 2020). In "Pedagogy of the Oppressed," Freire emphasizes the transformative power of education, driven significantly by the teacher's ability to guide and inspire. The importance of this role is magnified in an era where information is ubiquitous, making the teacher's guidance crucial in navigating the vast sea of data to distill what is true, relevant, and insightful.

The impact of a dedicated teacher can be profound and long-lasting, often becoming evident long after the student has left the classroom (Palmer, 1998). Palmer's "The Courage to Teach" explores this enduring influence, emphasizing how teachers shape lives in ways that are often immeasurable yet deeply significant.

In summary, the role of teachers as the bearers of the torch of knowledge encompasses far more than academic teaching. It is a multifaceted responsibility that entails inspiring, empowering, and guiding students. As society continues to progress, the necessity for passionate, committed, and knowledgeable teachers remains undiminished, a constant beacon in the journey of lifelong learning and personal development.



## **Nurturing individual strengths**

Every student possesses distinct characteristics, encompassing individual strengths and problems. Educators possess a high level of proficiency in discerning these specific characteristics. Educators focus on cultivating a student's inherent abilities, whether they lie in the realms of arts, sciences, athletics, or leadership. They also recognize areas of improvement and provide additional support, be it through extra classes, personalized assignments, or mentoring.

## **Instilling discipline and moral values**

Educators assume a pivotal role in molding the moral and ethical development of their students. Teachers establish discipline by implementing classroom laws, providing instructions for tasks, and facilitating interactive conversations. In addition, they facilitate the transmission of moral teachings, either by means of carefully selected curriculum subjects or by drawing from real-life instances, thus cultivating virtues such as truthfulness, uprightness, compassion, and tenacity (Lumpkin, 2008).

## **Emotional and psychosocial support**

For a significant number of students, the educational environment might present emotional difficulties that are equally as demanding as the cognitive challenges they face. The period of adolescence is characterized by significant challenges, as students frequently confront a wide array of concerns encompassing peer influence and struggles related to self-identity. In such situations, educators frequently assume the additional role of counsellors (Mazzer & Rickwood, 2015). Although robots have the capability to be programmed for the recognition and replication of human emotions, the qualities of real empathy, understanding, and emotional intelligence remain exclusive to humans. Educators provide a secure environment wherein students are able to openly articulate their apprehensions, hopes, and ambitions. The enduring emotional connection, which frequently endures for the entirety of one's life, is irreplaceable.

## **Preparing students for the future**

Beyond the confines of textbooks and exams, teachers prepare students for life. This includes soft skills like communication, teamwork, and critical thinking. Especially in higher grades, teachers also offer guidance regarding future career paths, helping students navigate the maze of higher education and job prospects.

## **Continuous evolution**

The role of a teacher is not static. The classroom setting undergoes transformation in tandem with global developments. Teachers have embraced technological advancements, such as the utilization of smart boards, the

integration of e-learning platforms, and the inclusion of social media in instructional practices. The global COVID-19 epidemic and the accompanying transition to remote learning have shown the remarkable capacity for adaptation demonstrated by educators around the globe.

## **Lifelong impact**

Perhaps the most profound role of a teacher is the lasting impact they leave on their students. Memories of school days often revolve around particular teachers – ones who believed in a student when no one else did, ones who offered words of wisdom, or ones who facilitated comprehension of challenging subjects by means of inventive pedagogical approaches.

## **Moral and ethical judgments**

Teaching is not just about transmitting information; it is about instilling values. Human educators can impart moral and ethical lessons, often drawing from personal experiences or cultural nuances. In essence, educators serve as the fundamental support system inside the framework of the educational system. The responsibilities they undertake extend beyond basic job titles. They exert influence on the future by impacting individual students in various manners. Understanding their role is crucial, not just for policymakers and educational institutions but for society at large. As the adage goes, "It takes a big heart to shape little minds," and teachers around the world do this with unparalleled dedication and passion. Therefore, despite the advancements in AI and robotics, it is unlikely that AI robots will completely replace teachers in the education system (Mubin et al., 2013).

## **The complementary roles of human and robot educators**

The increasing integration of AI and robots in educational settings has prompted inquiries concerning the potential responsibilities that these automated instructors may assume in conjunction with human educators. Nevertheless, instead of perceiving this phenomenon as a contest, it is advantageous to comprehend the ways in which human and robotic instructors may mutually enhance one another's abilities, so fostering a harmonious, efficient, and efficacious educational setting. In the study conducted by Huang et al. (2023), it is posited that human educators would continue to occupy a central role in the utilization of technical advancements such as digital twins, avatars, and physical robots.

Human educators have inherent characteristics that are fundamental to their being, including empathy, intuition, flexibility, and the capacity to establish authentic connections. Educators have a multifaceted role in the education system, since they not only impart information but also play a crucial role in fostering students' emotional and social development. When a student faces personal challenges, it is often the human touch, the reassuring word, or the understanding nod that makes a difference. Humans

can interpret subtle cues, from a student's tone of voice to their body language, gauging their comprehension, interest, or emotional state. Furthermore, they can inspire, mentor, and ignite passion, going beyond the curriculum to teach essential life lessons. On the other hand, robot educators bring to the table a suite of unique advantages that humans cannot match as previously noted. With vast data storage and processing abilities, they can offer personalized learning experiences tailored to each student's pace and learning style. Their unbiased nature ensures consistent instruction, free from any unintended biases. Robots can sift through vast amounts of data to provide real-time feedback and can make adaptive lesson plans based on a student's performance, ensuring that learning is always optimized. Additionally, their tireless nature means that they can be available round the clock, providing additional resources or assistance whenever a student may want it.

Consider the potential of using the respective advantages of both entities. In an integrated educational setting, the human instructor is primarily responsible for delivering instructional content, cultivating critical thinking skills, and addressing socio-emotional difficulties. Complementing this role, a robotic instructor can contribute by administering tailored assignments, promptly evaluating student submissions, and providing supplementary materials to support individuals requiring additional assistance. The integration of AI in the classroom facilitates enhanced efficiency by enabling educators to focus on more critical matters, as they can rely on AI to effectively manage administrative chores and personalized learning routes in the same way that Robot Abigail would be supporting her co-headteacher in Cottesmore School in the UK.

Moreover, robot educators can ensure that no student is left behind. In larger classrooms, where it is challenging for a single teacher to cater to the needs of every student, robots can provide the additional support required, ensuring that each student receives the attention they need. Conversely, human educators possess the ability to intervene when students require emotional assistance, direction, or motivation, which are domains where robots currently exhibit limitations.

However, as we embrace this combined approach, it is crucial to set boundaries and guidelines. The human aspect of education must never be overshadowed. Robots should be seen as tools that enhance the educational experience, not replace the intrinsic human elements. Their role should be to aid, not to lead.

The forthcoming trajectory of education does not hinge upon the dichotomy of selecting between human or robotic instructors, but rather on comprehending the potential synergy that may be achieved via their collaborative efforts. By integrating the cognitive qualities of individuals, such as emotional intelligence, flexibility, and mentoring skills, with the data-driven, consistent, and personalized methodologies employed by robots, we may provide a foundation for an enhanced, harmonious, and progressive educational setting. If appropriately cultivated, this mutually beneficial association has the potential to usher in a period characterized by deep and enhanced learning, effectively

harnessing the advantages offered by both realms.

## Conclusion

In the evolving landscape of education, characterized by the emergence of transformative technologies such as AI robots, it is imperative to acknowledge the vital role of human educators. Although AI has the ability to enhance and optimize certain activities, the fundamental nature of education, which is based on empathy, understanding, and authentic human interaction, remains unmatched. Teachers perform multiple responsibilities, from being knowledge facilitators to mentors and emotional anchors. Despite their numerous capabilities, robots are unable to completely imitate the intricate human touch that is crucial for comprehensive learning. In the current era characterized by rapid technology advancements, it is crucial to not only welcome novel advances but also recognize and support the enduring importance of human educators. Educators serve as more than just conveyors of information; they constitute the fundamental essence of the educational process, molding intellects and destinies in manners that are beyond the capabilities of machines.

The inquiry revolves around the potential for educators to include robots in their practices rather than the possibility of robots completely supplanting educators. Educators may secure their indispensability in the future educational environment by embracing technology and developing a comprehensive awareness of its capabilities and constraints. The fundamental essence of education perpetually resides within the realm of humanity. The proliferation of robots and artificial intelligence (AI) in the field of education represents a significant shift rather than a termination. Educators play a crucial role in navigating and embracing these changes since they possess the means to ensure that technology functions as a facilitator rather than a disruptor.

## References

- Alabdulkareem, A., Alhakbani, N., & Al-Nafjan, A. (2022). A systematic review of research on robot-assisted therapy for children with autism. *Sensors*, 22(3), 944. <https://doi.org/10.3390/s22030944>
- Aziz, A. A., & Ghanimi, M. H. A. (2020). Reading with robots: A personalized robot-based learning companion for solving cognitively demanding tasks. *International Journal on Advanced Science, Engineering and Information Technology*, 10(4), 1489-1496. <http://dx.doi.org/10.18517/ijaseit.10.4.7077>
- Cagiltay, B., White, N. T., Ibtasar, R., Mutlu, B., & Michaelis, J. (2022). Understanding factors that shape children's long term engagement with an in-home learning companion robot. In *Interaction design and children* (pp. 362-373). <https://doi.org/10.1145/3501712.3529747>
- Catlin, D., Kandlhofer, M., Holmquist, S., Csizmadia, A. P., Angel-Fernandez, J., & Cabibihan, J. J. (2018). EduRobot taxonomy and Papert's paradigm. In V. Dagiene & E. Jasute

- (Eds.), *Constructionism: Constructionism, computational thinking and educational innovation* (pp. 151–159). Vilnius, Lithuania.
- Conti, D., Cirasa, C., Nuovo, S. D., & Nuovo, A. D. (2020). Robot, tell me a tale! A social robot as tool for teachers in kindergarten. *Interaction Studies*, 21(2), 220–242. <https://doi.org/10.1075/is.18024.con>
- Edwards, B. I., & Cheok, A. D. (2018). Why not robot teachers: Artificial intelligence for addressing teacher shortage. *Applied Artificial Intelligence*, 32(4), 345–360. <https://doi.org/10.1080/08839514.2018.1464286>
- Freire, P. (2020). Pedagogy of the oppressed. In *Toward a sociology of education* (pp. 374–386). Routledge. <https://envs.ucsc.edu/internships/internship-readings/freire-pedagogy-of-the-oppressed.pdf>
- Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, 62, 729–754. <https://doi.org/10.1613/jair.1.11222>
- Greca Dufranc, I. M., García Terceño, E., Fridberg, M., Cronquist, B., & Redfors, A. (2020). Robotics and early-years STEM education: The botSTEM framework and activities. *European Journal of STEM Education*, 1, 1–13. <https://doi.org/10.20897/ejsteme/7948>
- Green, C. A., Mahuron, K. M., Harris, H. W., & O'Sullivan, P. S. (2019). Integrating robotic technology into resident training: Challenges and recommendations from the front lines. *Academic medicine: Journal of the Association of American Medical Colleges*, 94(10), 1532. <https://doi.org/10.1097/ACM.0000000000002751>
- Han, Z., Tu, Y., & Huang, C. (2023). A framework for constructing a technology-enhanced education metaverse: Learner engagement with human–machine collaboration. *IEEE Transactions on Learning Technologies* 16(6), 1179–1189. <https://doi.org/10.1097/10.1109/TLT.2023.3257511>
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Hooper, R. (2018). Get ready for the robot invasion of our classrooms. *The Japan Times*. <https://www.japantimes.co.jp/news/2018/08/28/national/science-health/get-ready-robot-invasion-classrooms/>
- Hu, Y. H., Fu, J. S., & Yeh, H. C. (2023). Developing an early-warning system through robotic process automation: Are intelligent tutoring robots as effective as human teachers? *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2022.2160467>
- Huang, R., Tlili, A., Xu, L., Chen, Y., Zheng, L., Metwally, A. H. S., Da, T., Chang, T. W., Wang, H., Mason, J., Stracke, C. M., Sampson, S., & Bonk, C. J. (2023). Educational futures of intelligent synergies between humans, digital twins, avatars, and robots - The iSTAR framework. *Journal of Applied Learning and Teaching*, 6(2), 28–43. <https://doi.org/10.37074/jalt.2023.6.2.33>
- Ifelebuegu, A. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning and Teaching*, 6(2), 385–392. <https://doi.org/10.37074/jalt.2023.6.2.2>
- Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in Education (AIEd) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching*, 6(2), 332–345. <https://doi.org/10.37074/jalt.2023.6.2.29>
- Johnson, S., & Acemoglu, D. (2023). Power and progress: Our thousand-year struggle over technology and prosperity. *International Affairs*, 99(6), 2525–2527. <https://doi.org/10.1093/ia/iad251>
- Joseph, O. U., Arikpo, I. M., Victor, O. S., Chidirim, N., Mbua, A. P., Ify, U. M., & Diwa, O. B. (2024). Artificial intelligence (AI) in academic research. A multi-group analysis of students' awareness and perceptions using gender and programme type. *Journal of Applied Learning and Teaching*, 7(1), 1–8. <https://doi.org/10.37074/jalt.2024.7.1.9>
- Kelly, J. (2023). *Goldman Sachs predicts 300 million jobs will be lost or degraded by artificial intelligence*. <https://www.forbes.com/sites/jackkelly/2023/03/31/goldman-sachs-predicts-300-million-jobs-will-be-lost-or-degraded-by-artificial-intelligence/?sh=3d590b7c782b>
- Lumpkin, A. (2008). Teachers as role models teaching character and moral virtues. *Journal of Physical Education, Recreation & Dance*, 79(2), 45–50. <https://doi.org/10.1080/07303084.2008.10598134>
- Mazzer, K. R., & Rickwood, D. J. (2015). Teachers' role breadth and perceived efficacy in supporting student mental health. *Advances in School Mental Health Promotion*, 8(1), 29–41. <https://doi.org/10.1080/1754730X.2014.978119>
- Mubin, O., Stevens, C. J., Shahid, S., Al Mahmud, A., & Dong, J.-J. (2013). A review of the applicability of robots in education. *Technology for Education and Learning*, 1, 1e7. <http://dx.doi.org/10.2316/Journal.209.2013.1.209-0015>
- Nourbakhsh, I. R. (2015). *Robot futures*. MIT Press.
- Palmer, P. J. (1998). *The courage to teach: Exploring the inner landscape of a teacher's life*. Jossey-Bass.
- Pandey, A. K., & Gelin, R. (2018). A mass-produced sociable humanoid robot: Pepper: The first machine of its kind. *IEEE Robotics & Automation Magazine*, 25(3), 40–48. <https://doi.org/10.1109/MRA.2018.2833157>
- Papadopoulos, I., Lazzarino, R., Miah, S., Weaver, T., Thomas, B., & Koulouglioti, C. (2020). A systematic review of the literature regarding socially assistive robots in pre-tertiary education. *Computers & Education*, 155, 103924. <https://doi.org/10.1016/j.compedu.2020.103924>
- Rosanda, V., & Istenic Starcic, A. (2019, September). The robot

in the classroom: A review of a robot role. In *International symposium on emerging technologies for education* (pp. 347-357). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-38778-5\\_38](https://doi.org/10.1007/978-3-030-38778-5_38)

Rudolph, J., Tan, S., & Aspland, T. (2023). Personal digital assistant or job killer? Generative AI and the teaching profession in higher education. *Journal of Applied Learning and Teaching*, 6(2), 07-16. <https://doi.org/10.37074/jalt.2023.6.2.1>

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. John Wiley & Sons. <https://www.wiley.com/en-gb/+Robots+Replace+Teachers%3F%3A+AI+and+the+Future+of+Education-p-9781509528967>

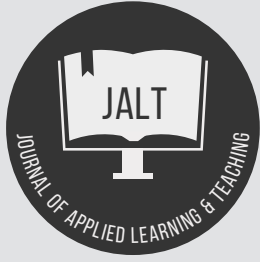
So, S., & Lee, N. (2023). Pedagogical exploration and technological development of a humanoid robotic system for teaching to and learning in young children. *Cogent Education*, 10(1), 2179181. <https://doi.org/10.1080/2331186X.2023.2179181>

Sperlinger, T. (2020). Aye, robot: Should robots replace teachers? *TLS: Times Literary Supplement*, (6108), 32-33. <https://www.the-tls.co.uk/articles/should-robots-replace-teachers-by-neil-selwyn-book-review/>

Tang, A. L., Tung, V. W. S., & Cheng, T. O. (2023). Teachers' perceptions of the potential use of educational robotics in management education. *Interactive Learning Environments*, 31(1), 313-324. <https://doi.org/10.1080/10494820.2020.1780269>

Tlili, A., Salha, S., Wang, H., & Huang, R. (2023). Intelligent tutoring systems examined in social experiments—Is the magic gone? A meta-analysis. In *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 50-54). <https://doi.org/10.1109/ICALT58122.2023.00020>.

Virgillito, M. E. (2017). Rise of the robots: Technology and the threat of a jobless future. *Labor History*, 58(2), 240-242. <https://doi.org/10.1080/0023656X.2016.1242716>



---

## Recognition of foreign professional degrees in Peru: Processes and strategies for improvement

---

Gianfranco Cavagnari-Bruce<sup>A</sup>      A      *Escuela de Posgrado. Universidad San Ignacio de Loyola. Lima, Perú*

Shyla Del-Aguila-Arcentales<sup>B</sup>      B      *Escuela de Posgrado. Universidad San Ignacio de Loyola. Lima, Perú*

Vladimir Rodriguez Cairo<sup>C</sup>      C      *Escuela de Posgrado. Universidad San Ignacio de Loyola. Lima, Perú*

DOI: <https://doi.org/10.37074/jalt.2024.7.1.13>

---

### Abstract

This brief article describes the recognition process of foreign professional degrees in Peru. A documentary review of the Peruvian regulations regarding the recognition process and its legal origin is carried out. Supreme Decrees, Laws and directorial resolutions were used as inputs. Flexibility was found in the recognition process of accepting foreign degrees in Peru. Recognizing foreign professional titles is based only on an administrative procedure and not on any academic criteria. An academic evaluation of professional qualifications is lacking in the recognition process as it is simply an administrative evaluation. The academic evaluation process could be based on a knowledge examination to ensure that the professionals to be recognized have a high level of knowledge. Several academic requirements must be incorporated to ensure entry into the labor market for suitable professionals of a high academic level.

**Keywords:** Education; Peru; professional degree; recognition of foreign professional degrees; universities.

### Introduction

The fall of the Berlin Wall is considered one of the most critical events in human history (Bohn & Berntsen, 2007; Marsh & Köller, 2004). Although globalization is an event that has existed for decades, the fall of the Berlin Wall has been an important event for the further expansion of human activities around the world (Hoskins & Todd, 2018; Sriramesh, 2008; Zurn, 2010). The wall's fall meant the collapse of European socialism, opening new markets from the West and motivating greater commercial exchange without borders (Monshipouri & Motameni, 2000; Vallet & David, 2012).

The growth of globalization has caused many benefits in the world's population, one of the effects of the increase in trade between countries (Borchert & Yotov, 2017; Deardorff & Stern, 2002; Harrison, 2006). Because of this, Europe opened all markets, making it easier for citizens to move from one country to another (Meunier, 2007). With the demographic movement, people also had to adapt to each country's new labor markets, customs and educational models. For the latter and other reasons, the Bologna Plan and the European Higher Education Area were created in the European Community (European Commission, 2023). The aim of this plan was to:

1. Introduce a three-cycle higher education system of bachelor's, master's and doctoral studies.
2. Ensure mutual recognition of qualifications and periods of learning abroad completed at other universities.
3. Implement a quality assurance system to strengthen the quality and relevance of learning and teaching.

Based on the European Union's Bologna Plan, several countries have recognized the need to update their procedures to register the studies of people who study in a different country. This becomes more urgent due to the greater immigration that exists in different regions, which leads professionals to work in other countries (Bastia & Piper, 2019; Segal, 2019).

Before the pandemic, global citizens' efforts had already been reported, which implies studying in different countries (Koukouraki, 2020). During COVID-19, virtual education allowed the internationalization of collaborations (Sam, 2022). After the pandemic, the migration of professionals

has increased further, boosting markets and education. However, there is a critical question: How is these professionals' training equivalent to or close to professionals in the destination country? When these migrants arrive in these countries, they seek to work in the professions using the qualifications they have obtained in their countries, so they need to register their professional titles in the National Registry of Degrees and Titles to practice their profession.

In this scenario, each country has a different mechanism to recognize professional degrees issued in other countries. Are they the most appropriate mechanisms? Can they be improved? This article presents the case of the procedures to register a professional title and demonstrates the evolution of recognition in Peru. This expected content can be considered for analysis by others to improve their professional title recognition processes.

## Methodology

A non-experimental, retrospective, qualitative, and descriptive design was used. The legal documents regulating professional degree recognition in Peru were evaluated. We searched the official Peruvian newspaper "El Peruano" web page. Likewise, the Ministry of Education, the Ministry of Foreign Affairs and the National Superintendence of University Higher Education (SUNEDU for its acronym in Spanish) are reviewed. The selection criteria for legal documents comprises current documents available on the institutions above' web pages. These documents must contain specific information on the processes of registration and recognition of professional degrees obtained in countries other than Peru. The main objective is to know the academic and administrative requirements requested by the Peruvian authorities to recognize professional degrees. The requirements requested by the National Superintendence of University Higher Education are evaluated.

## Results and discussion

Based on the information obtained, the processes of registration and recognition of professional degrees in Peru are described.

### Methods of registration of professional titles in Peru

In Peru, using Decree-Law No. 17437 - Organic Law of the Peruvian University (President of Peru, 1969a), the revalidation of titles, degrees and studies obtained or carried out abroad was implemented. According to the Decree Law, "revalidation could be carried out exclusively by specifically qualified universities" (President of Peru, 1969a). According to the Peruvian legal system, this would be the first norm in which there is a legal figure where a title, degree or study obtained abroad can be registered in Peru. Subsequently, in the same year, Decree Law No. 17662, which stipulates that professional degrees obtained in other countries is recognized without revalidation (President of Peru, 1969b), was published. In its first article, it is stated that professional degrees obtained in the universities of

countries with a treaty or agreement of cultural reciprocity shall be recognized without the requirement of revalidation. It also indicates that the institution in charge of registering them was the National Council of the University (President of Peru, 1969b).

The registration method for foreign professional degrees has not changed since 1969. In Peru, the National Superintendence of University Higher Education (SUNEDU) is the institution in charge of registering professional degrees and titles. There are two ways to register these professional degrees in the country: recognition and revalidation. Despite the current regulations, the revalidation and recognition processes continue in Peru. Considering that recognition is created as an exception to revalidation, we first proceed to explain what the latter consists of.

### About revalidation

According to Law 30220 - University Law (Congress of Peru, 2014), the National Superintendence of University Higher Education has as one of its functions "to establish the technical criteria for the validation and revalidation of studies, degrees and titles obtained in other countries." For this reason, the National Superintendence of University Higher Education Board of Directors, through Resolution No. 119-2019-SUNEDU/CD, establishes the technical criteria for the revalidation of degrees and titles granted abroad (National Superintendence of Higher University Education, 2019). In the National Superintendence of University Higher Education Board of Directors Resolution, revalidation is defined as:

...the procedure through which effects are granted in national territory to the degree or title granted by a foreign educational institution as a consequence of having passed an academic evaluation carried out by a university licensed to provide higher university educational service in Peru (National Superintendence of Higher University Education, 2019).

The technical criteria established in the resolution are presented in Figure 1. The Institutional Licensing criterion refers to the fact that to revalidate, the university must have a valid institutional license granted by the National Superintendence of University Higher Education (National Superintendence of Higher University Education, 2019). In the same sense, the criterion of academic experience refers to the degree program having at least ten graduating classes. The third criterion refers to the Academic Committee, which must have at least 3 members. At least one must be a regular professor, another a research professor, and all members must have a master's or doctorate in the same professional field related to the degree or title to be revalidated. Finally, they must comply with the other requirements of the University Law.

The fourth criterion refers to the evaluation of the revalidation, where if there is a minimum of 80% equivalence of the content of the courses, the revalidation can proceed. The university may choose the best way to reach 80% of similarity if the percentage is lower. The fifth criterion

indicates that if the degree has been obtained through research work or thesis; this must also have been presented. The sixth criterion is that the university must verify that the person performing the revalidation must also know a second language as per the University Law.

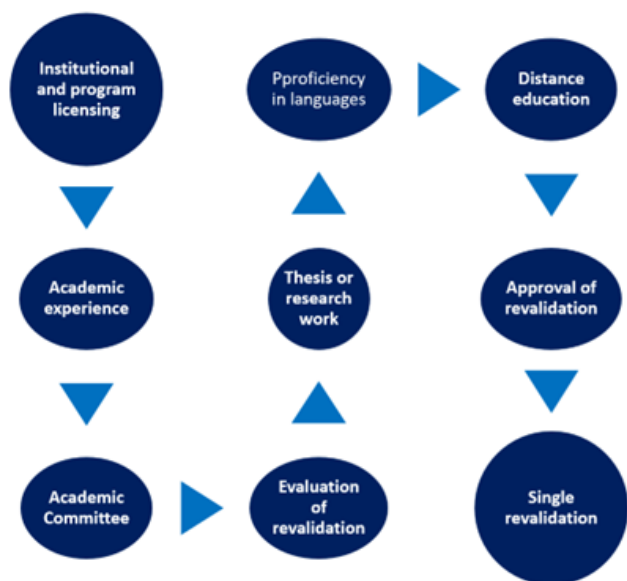


Figure 1. Enrollment in the National Registry of Degrees and Titles.

It has been detailed that recognition is granted by the National Superintendence of University Higher Education through administrative procedures when there are Instruments for the Recognition of Academic Degrees and Titles (IRGAT). Revalidation, which is the ordinary procedure in the absence of IRGAT, is a more complex procedure. In addition to the usual administrative procedures, revalidation involves the completion of an academic evaluation (Rojas, 2016). In the case of distance studies, the seventh criterion indicates that what is mentioned in the University Law is considered.

### About recognition

National Superintendence of University Higher Education, in its Resolution of the Board of Directors N° 099-2020-SUNEDU/CD (National Superintendence of Higher University Education, 2020) Regulations for the Recognition of Degrees and Degrees Granted Abroad, defines Recognition as:

...the administrative act by which the State, through National Superintendence of University Higher Education, grants effectiveness to the degrees and titles obtained abroad by evaluating their conformity with the quality criteria or the legal obligations assumed by the Republic of Peru by a treaty.

Likewise, it can be observed that among the requirements requested for the administrative process are not certificates of studies or academic transcripts. Only the diploma with the Apostille of The Hague is requested, or in any case, a database that corroborates the student's study (National Superintendence of Higher University Education, 2023) (Table 1).

Table 1. Comparison of criteria/requirements among recognition and revalidation.

Criteria/Requirements	Recognition	Revalidation
Diploma	Yes	Yes
Certificate of studies	No	Yes
Type of evaluation	Administrative	Academic
Proceed to	Agreements and Schedules	Programs of the university to be revalidated
Deadline	30 working days	According to university regulations

How could the National Superintendence of University Higher Education safeguard educational and professional quality in Peru, without information on the content of the courses it recognizes?

On the other hand, it has been reported that:

... by the principle of presumption of veracity, the recognition procedure admits the presentation of documents issued by foreign institutions or entities, with which there is not necessarily interoperability that would allow the National Superintendence of University Higher Education to verify their authenticity within the maximum legal term (Orellana, 2020, p. 23).

Reference is made to the former National Assembly of Rectors (ANR for its acronym in Spanish), and we compare the recognition procedure of the National Assembly of Rectors and National Superintendence of University Higher Education. In that case, it is observed that the level of evaluation has decreased. The recognition procedure at the National Assembly of Rectors was regulated by Ministerial Resolution No. 0035-2011-ED – Procedures for the recognition and total legal value in Peru of degrees granted by accredited universities by the States with which trade agreements, cultural agreements and others are signed. The Ministry of Education, through the Resolution mentioned above, defines recognition as the "act intended to grant academic or professional validity in Peru to an academic degree, professional title or certificates of studies obtained in a foreign university institution ..." (Ministry of Education of Peru, 2011).

As can be seen, the recognition of the National Assembly of Rectors also granted validity to degrees or titles. However, when reviewing the requirements requested by the National Assembly of Rectors, it can be observed that in addition to the National Superintendence of University Higher Education, the Certificate of Studies was requested. Furthermore, to recognize a masters or doctorate, the previous degrees were requested to be registered (Ministry of Education of Peru, 2011) (Table 2).

Table 2. Comparison of criteria among National Superintendence of University Higher Education and National Assembly of Rectors.

Criteria	National Superintendence of University Higher Education	National Assembly of Rectors
Diploma	Yes	Yes
Certificates of studies	No	Yes
Official translations	No	Yes
Simple translation	Yes	No

In addition, extra requirements were indicated when a degree in the medical field was recognized per Resolution No. 610-2008-ANR - Regulations of the Office of Recognition and Certificates of the General Secretariat of the National Assembly of Rectors.

The additional requirements for the medical area were proof of completion of the hospital rotation. If the internship was done abroad, it must be for six months (specialties of Medicine, Surgery, Gynecology-Obstetrics and Pediatrics). If the rotation was not performed in Peru, it must be for one year and in the same specialties as the foreigners. Finally, when the applicant has more than 3 years of professional experience, they may undergo an interview to recognize the applicant's experience (National Assembly of Rectors, 2008). Table 3 compares the requirements for recognizing foreign professional degrees in health between the National Superintendence of University Higher Education and the National Assembly of Rectors.

Table 3. Comparison of criteria among National Superintendence of University Higher Education and National Assembly of Rectors for foreign professional degrees.

Criteria	National Superintendence of University Higher Education	National Assembly of Rectors
Diploma	Yes	Yes
Certificates of Studies	No	Yes
Study plan	No	Yes
Hospital Rotation	No	Yes

The National Superintendency of Higher University Education (SUNEDU), in an official letter No. 9328-2022-SUNEDU-03-08-04, confirmed the number of approvals of professional titles in 2020 (2809), 2021 (4122) and January-April 2022 (1521).

### Opportunities for improvement

The system can be improved through a crowdsourcing strategy to ensure the academic-professional level of the professionals working in Peru. In this regard, the professional associations should be convened to analyze the previous and current processes, evaluate the positive and negative aspects of both processes, and develop a consensus proposal to improve the process, ensuring a comprehensive evaluation of the academic and professional quality of the applicant.

On the other hand, it is necessary to evaluate the requirements requested in other countries to recognize, through benchmarking, the most beneficial aspects that can be incorporated into the current process in Peru. The information provided by countries in Latin America, such as Argentina (Ministry of Education of Argentina, 2023), Bolivia (Ministry of Education of Bolivia, 2023), and Chile (Ministry of Foreign Affairs of Chile, 2023), can be taken as an example. Information from Oceania (Medical Board Ahpra, 2023), Africa (South African Qualifications Authority, 2023), Asia (Ministry of Labor of Taiwan, 2023), and Europe (Federal Ministry of Education and Research of Germany, 2023) is valuable to consider in generating a proposal to improve the process of recognition of professional qualifications.

### Conclusion

An analysis of the procedures for registering a professional degree in Peru shows that recognition is merely an administrative procedure without academic review; on the other hand, revalidation is the ordinary procedure for registering a foreign professional degree in Peru, and when there is a trade agreement, recognition can be made. When comparing the requirements demanded by the former National Assembly of Rectors and the current National Superintendence of University Higher Education, it is recognized that the requirements have been decreasing and, consequently, the evaluation level.

Regarding professional medical degrees, the former National Assembly of Rectors requested more requirements such as "hospital rotation". This decrease in the requirements can generate the situation that some professionals do not have an academic-professional level equivalent to Peruvian professionals, with the related risks and consequences, especially in medicine. In this sense, it can be concluded that the recognition process in Peru has been losing its requirements and quality, which may allow a professional who does not have the specific knowledge of Peru to practice as a professional.

### References

Bastia, T., & Piper, N. (2019). Women migrants in the global economy: A global overview (and regional perspectives). *Gender & Development, 27*(1), 15-30. <https://doi.org/10.1080/13552074.2019.1570734>

Bohn, A., & Berntsen, D. (2007). Pleasantness bias in flashbulb memories: Positive and negative flashbulb memories of the fall of the Berlin Wall among East and West Germans. *Memory & Cognition, 35*(3), 565-577. <https://doi.org/10.3758/BF03193295>

Borchert, I., & Yotov, Y. V. (2017). Distance, globalization, and international trade. *Economics Letters, 153*, 32-38. <https://doi.org/10.1016/j.econlet.2017.01.023>

Congress of Peru. (2014). *Law 30220 - University Law of 2014*. <https://diariooficial.elperuano.pe/pdf/0021/ley-universitaria-30220.pdf>

Deardorff, A. V., & Stern, R. M. (2002). What you should know about globalization and the world trade organization. *Review of International Economics, 10*(3), 404-423. <https://doi.org/10.1111/1467-9396.00340>

European Commission. (2023). *Bologna process and European higher education area*. <https://education.ec.europa.eu/es/education-levels/higher-education/inclusive-and-connected-higher-education/bologna-process>

Federal Ministry of Education and Research of Germany. (2023). *Information on professional recognition*. <https://www.anerkennung-in-deutschland.de/html/en/index.php>

Harrison, A. (2006). *Globalization and poverty*. <https://www.>



Hoskins, W. L., & Todd, W. F. (2018). Twenty years after the fall of the Berlin Wall: Rethinking the role of money and markets in the global economy. *Levy Economics Institute, Working Papers Series* (908). <http://dx.doi.org/10.2139/ssrn.3189059>

Koukouraki, K. (2020). Supporting students in developing critical global citizenship: Examples from the English for Academic Purposes (EAP) classrooms. *Journal of Applied Learning and Teaching*, 3(2), 107-115. <https://doi.org/10.37074/jalt.2020.3.2.16>

Marsh, H. W., & Köller, O. (2004). Unification of theoretical models of academic self-concept/achievement relations: Reunification of east and west German school systems after the fall of the Berlin Wall. *Contemporary Educational Psychology*, 29(3), 264-282. [https://doi.org/10.1016/S0361-476X\(03\)00034-1](https://doi.org/10.1016/S0361-476X(03)00034-1)

Medical Board Ahpra. (2023). *Registration*. <https://www.medicalboard.gov.au/Registration.aspx>

Meunier, S. (2007). Managing globalization? The EU in international trade negotiations. *JCMS: Journal of Common Market Studies*, 45(4), 905-926. <https://doi.org/10.1111/j.1468-5965.2007.00753.x>

Ministry of Education of Argentina. (2023). *Validate your foreign university degree*. <https://www.argentina.gob.ar/educacion/tramites/convalidaciones-universitarias-extranjeros>

Ministry of Education of Bolivia. (2023). *Revalidation and homologation of professional titles*. [https://www.minedu.gob.bo/files/documentos-normativos/VESFP/utp/requisitos/Revalidacion\\_Homologacion.pdf](https://www.minedu.gob.bo/files/documentos-normativos/VESFP/utp/requisitos/Revalidacion_Homologacion.pdf)

Ministry of Education of Peru. (2011). *Ministerial Resolution N°35-2011 procedure for the recognition and full legal value in Peru of degrees awarded by accredited universities by states with which trade agreements, cultural and other agreements have been signed*. <https://www.gob.pe/institucion/minedu/normas-legales/166382-0035-2011-ed>

Ministry of Foreign Affairs of Chile. (2023). *Recognition of academic degrees for foreigners*. <https://www.chile.gob.cl/bogota/tramites/para-extranjeros/reconocimiento-de-titulos-academicos-para-extranjeros>

Ministry of Labor of Taiwan. (2023). *Documents required for foreign professional employment work permit application*. <https://ezworktaiwan.wda.gov.tw/en/cp.aspx?n=8E87472D9CB255FF&s=BCA6B6757A58F5B3>

Monshipouri, M., & Motameni, R. (2000). Globalization, sacred beliefs, and defiance: Is human rights discourse relevant in the Muslim world? *Journal of Church and State*, 42(4), 709-736. <https://doi.org/10.1093/jcs/42.4.709>

National Assembly of Rectors. (2008). *Resolution 610-2008-ANR - regulations of the recognition and certifications office*. [https://cdn.www.gob.pe/uploads/document/file/765716/resolucion\\_consejo\\_di...20200603-2874-1s153se.pdf](https://cdn.www.gob.pe/uploads/document/file/765716/resolucion_consejo_di...20200603-2874-1s153se.pdf)

National Superintendency of Higher University Education. (2019). *Technical criteria for the revalidation of degrees and titles awarded abroad*. <https://sunedu.gob.pe/resolucion-del-consejo-directivo-n119-2019-sunedu-cd/>

National Superintendency of Higher University Education. (2020). *Regulations for the recognition of degrees and/or diplomas awarded abroad*. <https://www.sunedu.gob.pe/resolucion-del-consejo-directivo-no-099-2020-sunedu-cd>

National Superintendency of Higher University Education. (2023). *Procedure for recognition of degrees and titles*. <https://www.sunedu.gob.pe/procedimiento-de-reconocimiento-de-grados-y-titulos-extranjeros/>

Orellana, C. L. K. (2020). *Implications of administrative simplification in the procedure for the recognition of degrees and diplomas awarded abroad*. PUCP. <https://tesis.pucp.edu.pe/repositorio/handle/20.500.12404/16327>

President of Peru. (1969a). *Organic law of the Peruvian university*. <https://docs.peru.justia.com/federales/decretos-leyes/17437-feb-18-1969.pdf>

President of Peru. (1969b). *Professional degrees obtained in other countries will be recognized without revalidation*. <https://docs.peru.justia.com/federales/decretos-leyes/17662-may-27-1969.pdf>

Rojas, E. J. (2016). *The instruments of recognition of academic degrees and professional titles as a tool of Peru's foreign policy: Balance and proposals* [Master, Academia Diplomática del Perú]. Lima - Peru. <https://core.ac.uk/reader/250085280>

Sam, C. Y. (2022). Post-COVID-19 and higher education. *Journal of Applied Learning and Teaching*, 5(1), 156-164. <https://doi.org/10.37074/jalt.2022.5.1.21>

Segal, U. A. (2019). Globalization, migration, and ethnicity. *Public Health*, 172, 135-142. <https://doi.org/10.1016/j.puhe.2019.04.011>

South African Qualifications Authority. (2023). *Recognition of professional bodies*. <https://www.sqaq.org.za/services/recognition-of-professional-bodies/>

Sriramesh, K. (2008). Globalization and public relations. In A. Zerfass, B. van Ruler, & K. Sriramesh (Eds.), *Public relations research: European and international perspectives and innovations* (pp. 409-425). VS Verlag für Sozialwissenschaften. [https://doi.org/10.1007/978-3-531-90918-9\\_27](https://doi.org/10.1007/978-3-531-90918-9_27)

Vallet, É., & David, C-P. (2012). Introduction: The (re)building of the wall in international relations. *Journal of Borderlands Studies*, 27(2), 111-119. <https://doi.org/10.1080/08865655.2012.687211>

Zurn, M. (2010). Fall of the Berlin Wall: Globalisation and the future of Europe. *New Zealand International Review*, 35(3), 2-7.

Copyright: © 2024. Gianfranco Cavagnari-Bruce, Shyla Del-Aguila-Arcetales and Vladimir Rodriguez Cairo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



These chapters include:

1. What does it mean to teach well?
2. How do our experiences as students frame our teachings?
3. How do we do learning-centred teaching?
4. How do we deal with classroom and self-sabotage?
5. How does power show up in classrooms?
6. How do we democratise classrooms?
7. How do we promote good discussions?
8. Teaching critical thinking
9. How do we teach about race?
10. How can we become critically reflective teachers?
11. How can we enact the power of modelling?
12. How do we teach well as leaders?
13. How do we learn and grow as teachers over a career?

As Rudolph and Tan explain, the questions that became the themes of the chapters are also those that have repeatedly surfaced at the hundreds of teacher development workshops that Brookfield has conducted over his long and illustrious career. It is little wonder, then, that what emerges in the narrative are long-considered and studied responses, thick with Brookfield's personal history and the humility with which he brings to his subject.

What is especially striking about the book is the portrait that emerges of Brookfield himself and the way in which this provides a kind of hopeful reassurance to the intended audience. Here is an internationally renowned scholar who has dedicated his life to teaching, and yet he candidly reveals his struggles as a student, his insecurities and the strong 'imposter syndrome' he has experienced. Unquestionably, this will resonate with many teachers and encourage those teachers who may be experiencing similar anxieties and self-doubt. Indeed, Brookfield notes that these personal struggles are a useful resource for teachers as they better enable them to understand and help students with similar challenges.

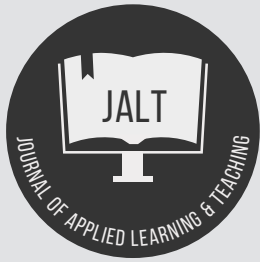
What I really value about *'Teaching well'* is the way in which it addresses what might be considered the subtle nuances of classroom dynamics. For instance, there is a thoughtful consideration of 'how it sounds to teach well', the critical importance of context and, importantly, how good teachers make many procedural decisions to account for variations in context. Additionally, the book contains a wealth of practical tips and practices for managing classroom dynamics, including how to identify student learning challenges, deal with sensitive discussion topics, anonymous feedback mechanisms, and promote productive discussions on how to address 'sabotage' in a classroom setting.

A particular strength of the book is the way in which the conversations touch on and discuss key theoretical concepts – for example, Foucault's contributions to our understanding of power are discussed in relation to the way in which power 'shows up' in classrooms. This fusion of the theoretical with the lived experience of teaching is helpful. For readers who may not be familiar with Foucault's work, consumption of the complex ideas expressed may prove challenging. On the other hand, these readers may be inspired to read more of Foucault and some of the other works of the venerated theoreticians that are mentioned throughout the book.

Following the work and life story of Professor Brookfield inevitably means not shying away from the more difficult and potentially controversial topics. There is, for example, a deep exploration of the issue of race and racism in higher education, which moves from how we might teach about race to the silencing of Critical Race Theory and managing racism in the classroom.

There is much to be admired about *'Teaching well'*, from the accessibility of its conversationally-driven narrative to the juxtaposition of practical teaching ideas with thoughtful consideration of established pedagogical theory. For teachers at any level of education, the book offers unique access to the considerable wisdom of Professor Brookfield on a broad range of topics skilfully presented by Rudolph and Tan. Teachers will also discover, I'm sure (as I did), something more than just a guidebook for teaching well. The book additionally offers a compass for how teachers themselves might wish to develop and manage their careers as critically reflective professionals. *'Teaching well'* is, therefore, a book that every teacher would do well to read and have close to them as they progress in their careers.

Copyright: © 2024. Peter Waring. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

Book review of Lindgren, Simon (Ed., 2023). Handbook of critical studies of artificial intelligence. Edward Elgar.

Jürgen Rudolph<sup>A</sup>

A

*Director of Research & Learning Innovation, Kaplan Higher Education Academy*

DOI: <https://doi.org/10.37074/jalt.2024.7.1.30>

## Introduction

Due to its hard cover, size and weight, this gigantic 940-page Handbook can be used literally as a weapon. The 76 chapters, contributed by 127 authors, also serve as a metaphorical weapon against the irrational exuberance surrounding AI, a phenomenon also observed in higher education (Rudolph et al., 2024). While the book is not focused on AI and higher education, it is an important book on the critically important topic of critical AI studies (pun intended) that it is highly suitable for a review in JALT.

Receiving this impressive tome was like finding myself at the doorstep of the fabled gingerbread house in the Brothers Grimm's haunting tale of Hänsel and Gretel – lost in a huge, foreboding forest, my breadcrumb trail disappeared, standing before a cottage built from confections, after having been abandoned by my parents. This book, much like the witch's cottage, is laden with revelations about the darker sides of AI, ready to devour any naive preconceptions we hold about this technology. Stretching this metaphor a tad further, the volume, however, serves not as a trap but as a beacon of critical theory, challenging us to confront and overthrow the cannibalistic tendencies of witchy Big Tech. In this analogy, the book embodies Gretel's cunning courage, enabling us to shove the menacing forces of unchecked technological advance into the oven, incinerating its malevolent underpinnings and illuminating a path toward a more critical understanding of AI.

Impressively, the Handbook has a single editor, Simon Lindgren, a Professor of Sociology and a Director of the DIGSUM Centre for Digital Social Research at Umeå University, Sweden. Lindgren's Handbook gathers cutting-edge insights from scholars across a wide variety of disciplines with a sense of urgency. It critically examines AI's expanding influence in society and culture and broadens the discussion beyond mere technological aspects by thematising AI's social, ethical, and political impacts. Covering key issues such as biases within AI systems, effects on democracy, privacy concerns, and its role in decision-making, the book advocates for a rigorous critique of AI. Lindgren's Handbook significantly succeeds in enriching the debate on critical AI

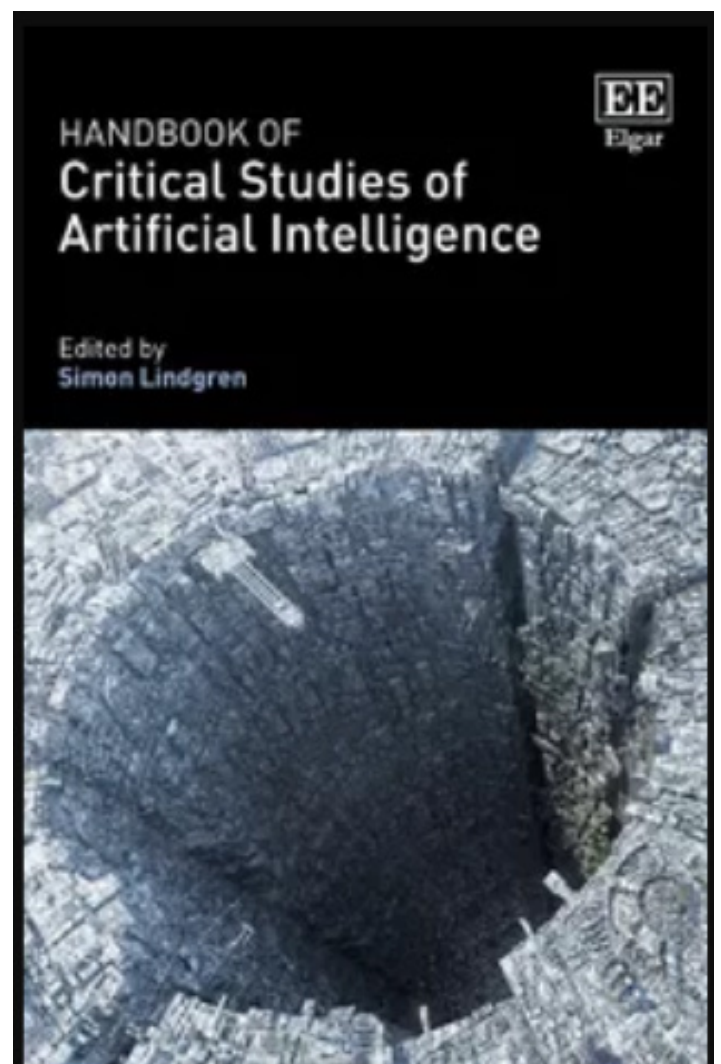


Figure 1: Book cover of Lindgren (Ed., 2023).

studies, posing significant questions, and introducing vital concepts for analysis.

In Lindberg's excellent introductory chapter, he rightly highlights that much of the work that has been done on AI is not critical enough and "critical perspectives of the technology are more urgent than ever" (p. 1).

The rationale behind this handbook is that we need to push analyses of AI much further into critical territory than what is the case today. The handbook wants to contribute to an ongoing discussion about what critical studies of AI can entail – what questions it may pose, and what concepts it can offer to address them (p. 4).

At present, we are surrounded by so much hype and mania based on techno-optimism and solutionism (Rudolph et al., 2024). We experience “an almost perverse obsession with artificial intelligence” (Kotasek, p. 254). AI promises to revolutionise “science, automatically detect various health issues, eliminate hate speech and misinformation, and prevent crimes” (Jobin & Katzenbach, p. 43). AI is also seen “as the solution to society’s problems, including climate change, pandemics and the energy crisis” (Verdegem, p. 302). Brevini (Chapter 75) points out that AI technologies have become such integral parts of our daily routines that they often go unnoticed, permeating every socially, politically, and economically significant sector. Examples include AI-driven traffic management cameras, facial recognition systems at airports, smartphone apps that suggest music videos on YouTube, and smart homes operated by Amazon’s Alexa. In the Handbook’s critical perspectives, Lindgren and his army of authors are influenced by the critical theory of the Frankfurt School. Horkheimer (1972) conceptualised critical theory as the radical analysis of present socio-economic conditions so that the research can be a liberating influence with an emancipatory agenda to create a world which satisfies the needs of humans.

In his introductory chapter, Lindgren makes many excellent observations about AI. One of the starting points is Crawford’s (2021, p. 8) succinct observation that “AI is neither artificial nor intelligent” and that “AI is politics by other means”. It is of critical importance to highlight that ‘artificial intelligence’ is a highly problematic and loaded concept despite its seemingly magical advantages. AI is a marketing label vulnerable to exploitation and exaggeration, and Lindgren has posited elsewhere that “AI is driven by myths that animate it as magic” (2024, p. 94). In his introduction, Lindgren (p. 17) further states that AI, being “the subject of evolving wars of definitions”, is an “*empty signifier*”, a “*ubiquitous apparatus... entangled with human experience*”, and “part of the *technological unconscious*”. AI is an ideology-driven, socio-political mythology, “constantly repeated and performed in marketing talk, hyped-up conferences, tech evangelism, business manifestos, and overblown media reporting” (p. 17).

The tome is organised into seven parts: (1) AI and critical theory: Conceptual discussions, (2) AI imaginaries and discourses, (3) the political economy of AI: datafication and surveillance, (4) AI transparency: ethics and regulation, (5) AI bias, normativity and discrimination, (6) politics and activism in AI, and (7) AI and automation in society. Unfortunately, space limitations render it impossible to do justice to every single chapter of the book. I will provide an overview of all of its seven parts, cherry-pick some of my favourite chapters (of which there are many), and paint in broad strokes before concluding with a critical appraisal.

## AI and critical theory: Conceptual discussions

The first part of the Handbook focuses on conceptual issues related to critically analysing AI through academic research. 15 chapters dive into conceptual issues surrounding AI, exploring its relationship with politics, ideology, governance, and a host of other concepts like antagonism, epistemology, and (de)coloniality. They critically examine AI’s integration into society, its governance, and the ideological underpinnings it perpetuates within the capitalist framework. Overall, the Handbook’s first section urges a re-evaluation of many of the AI-related conceptual dimensions.

Bloom’s chapter (3) laments that current AI regulations leave “Big Tech virtually unscathed”, as they lack “a focus on those affected by AI systems, apparently missing any general requirement to inform people who are subjected to algorithmic assessment” (MacCarthy & Propp, 2021, cited in p. 33). This chapter also contains some thought-provoking perspectives on 21st-century robots as slaves. Bloom cites Hampton (2015, p. 2):

Slavery... was largely invested in producing and controlling a labor force, which was disassociated from humanity. In many regards, American slavery was a failed experiment to employ flesh and blood machines as household appliance[s], farm equipment, sex toys, and various tools of industry without the benefit of human and civil rights... The technology of the 21st century is in the process of developing a modern-day socially accepted slave (cited on p. 39).

Chiodo’s chapter (6) contains a fabulous passage that shows how AI usurps divine dimensions by being “immanently omnipresent, omniscient, omnipotent and inscrutable”: “omnipresence (by being everywhere: it is always with us), omniscience (by knowing everything, from the answers to our questions to ourselves by tracking us), omnipotence (by increasingly having power over us, from shaping our worldviews to shaping our decisions and actions accordingly) and inscrutability (algorithms as black boxes)” (p. 76).

Birhane and Talat conclude their chapter (11) by showcasing the Te Hiku NLP project as an instance of decolonial AI. This initiative involved vast numbers of Te Reo Māori speakers in Aotearoa/New Zealand engaging collectively in the creation of AI technologies aimed at revitalising their endangered language. Over a span of ten days, the Te Hiku NLP project gathered 310 hours of spoken language and corresponding text from 200,000 recordings by 2,500 participants. This extensive dataset enabled the development of a speech recognition model with an 86% accuracy rate, driven by a desire to safeguard Māori culture and language. The success of the Te Hiku project in Te Reo Māori demonstrates the potential for machine learning systems to contribute to decolonisation efforts.

## AI imaginaries and discourses

AI is not only a technology, but also a story. (Coeckelbergh, 2021, p. 1626)

The Handbook's second part addresses the social and symbolic framing of AI and its connection to ideology. At present, AI occupies a central role in a broad spectrum of social and political imaginings, ranging from utopian dreams to dystopian fears. These imaginings typically emphasise notions of progress, profit, precision, and rationality while grappling with fears of robot-led catastrophes and disenchantment. AI shapes our perception of reality and the kinds of future lives and societies we deem achievable. Lagerkvist (2020, p. 16) states that "In the present age, AI emerges as both a medium for and a message about (or even from) the future, eclipsing all other possible prospects" (cited on p. 7).

The "AI imaginaries and discourses" section consists of nine chapters and is propelled by ideas often tied to societal power dynamics, which are usually portrayed as unbiased truths and objectives. These ideas commonly include faith in technology's salvational potential and the propensity to defer crucial decisions to inscrutable automated systems. The risk lies in accepting these views and priorities as infallible. The alternative proposed is a shift towards democratic decision-making and a thorough critique of AI's societal impact.

In Chapter 19, Kajava and Sawhney advise against the common practice in popular science and culture of attributing exaggerated human-like qualities to AI, such as *learning*, *training*, or *memory*, and using hyperbole. They highlight that AI technologies are often inaccurately described as possessing human-like agency or autonomy, a misconception stemming from equating machine *behaviour* with human action (Searle, 1992). Kajava and Sawhney note the confusion arising from attributing autonomy, agency, or sentience to machine behaviour. Following Rehak (2021), they suggest eliminating the use of terms such as 'agency' and 'autonomy' in discussions about AI due to their potential for inaccuracy and misleading implications. Instead, they call to refocus attention on the human responsibility inherent in AI's creation, deployment, and oversight.

In Chapter 20, Ballatore and Natale critique the conventional 'rise and fall' narrative used to describe AI's history, where periods of enthusiasm ('summers') alternate with times of disappointment ('winters'). They suggest a reinterpretation that consistently recognises the role of controversy and scepticism in shaping AI's development rather than viewing it as a sequence of distinct optimistic or pessimistic phases. This perspective reveals that doubts and debates have persistently accompanied AI from its inception in the 1950s to today, challenging the simplistic narrative of technological progress. They advocate for a critical examination of AI that considers failures and controversies as central to understanding its evolution. This approach serves as a corrective to current discussions about AI, underlining both its potential and its limitations, and attacks the long-held myth of machines capable of thinking like humans. It is worth remembering that Alan Turing (1950) considered the question 'Can machines think' "too meaningless to deserve discussion".

Verdicchio's chapter (21) makes a compelling distinction between 'artificial' and human intelligence and challenges the idea that intelligence can be fully captured in precise,

machine-compatible descriptions. He supports Gardner's (1983) concept of multiple intelligences, which categorises human intelligence into eight distinct types: visual-spatial, linguistic-verbal, logical-mathematical, bodily-kinaesthetic, musical, interpersonal, intrapersonal, and naturalistic. Verdicchio points out that in the pursuit of Artificial General Intelligence (AGI), the objective is highly ambitious: to translate every task humans are capable of – encompassing Gardner's entire spectrum of intelligences – into computational terms that machines can perform.

I would be remiss not to briefly mention Lina Rahm's excellent chapter (25) on "Educational imaginaries of AI". Rahm's chapter explores and questions the construction of knowledge regarding AI within international education and AI ethics policies and examines the social, political, and epistemic implications of this knowledge. Her chapter's "purpose is to support the development of critically reflexive and just education policies about AI futures" (p. 289).

### **The political economy of AI: Datafication and surveillance**

[D]omination perpetuates and extends itself not only through technology but as technology, and the latter provides the great legitimization of the expanding political power, which absorbs all spheres of culture. (Herbert Marcuse, cited in Timcke, p. 323)

The Handbook's third part contains eight chapters that explore the political economy of AI, particularly how AI and automation, in their deployment and outcomes, intersect with societal structures of dominance and exploitation, especially regarding capital and labour. Lindgren posits that critical examinations of AI should consider the social, political, and economic contexts surrounding the technology, along with its effects on these areas. It is essential to recognise that technology is inherently intertwined with the political economy. Consequently, it should not be viewed as independent from society.

Timcke (Chapter 28) reminds us that at the close of the last century, there was widespread optimism among various groups, from Habermasian scholars enthusiastic about the potential for a rejuvenated public sphere to techno-libertarians eager to explore new forms of non-state governance. This optimism was rooted in the digital revolution, with beliefs that the adoption of free software and a shift away from private property rights would lead to a new era of *homesteading on the electronic frontier* (Rheingold, 1994). However, reality has unfolded differently: capitalist agendas have increasingly restricted and dominated digital spaces. Timcke turns to the seldom-invoked Marxist concept of reification to argue that the deepening of capitalist social relations has significantly narrowed the possibilities for imagining life beyond these confines. This intensification of capitalist ideology has not only strengthened its influence but also fostered its acceptance among workers themselves, suggesting a profound shift in societal reasoning towards capitalist logic.

## AI transparency, ethics and regulation

Technological progress cannot count as progress without social evaluations of what it is good for, who benefits from it, and what costs it generates... progress is a reflexive concept; every progressive process must be constantly questioned as to whether it is in the social interest – correctly understood – of those who are part of this process. Thus, every criticism is itself also part of progress. (Forst, 2017, pp. 72-73, cited on p. 400)

The Handbook's fourth part focuses on ethics and regulation. In his introductory chapter, Lindgren mentions that researchers in critical AI studies might view areas such as "ethical AI" and "responsible AI" as insufficiently critical. However, critical theory faces a significant challenge in enhancing AI ethics rather than simply critiquing it. There certainly is much to be critical about. For instance, in Chapter 47, Albert and Delano state that much research on AI ethics is "shallow and unspecific, focusing on hypothetical or abstract potential harms rather than the specific impacts on particular groups" (p. 538).

Resseguier (Chapter 35) emphasises the lasting value of ethical ideals such as dignity, fairness, privacy and transparency. The best way to bring about these ideals may be to recognise the 'nonideal'. The near-impossible challenge is to ensure that these high-level principles are implemented for all.

Lee et al. (Chapter 36) remind us that the AI arms race – the competitive desire to create more powerful AI – facilitates harms that disproportionately impact marginalised groups by encouraging the development of insufficiently vetted AI systems. AI regulation faces fragmentation due to competitive dynamics, the push for innovation, the varied applications of AI, and its technical complexities. This means that creating uniform regulations would be impractical. Lee et al. argue that regulation must be adaptive and acknowledge divergent perspectives, social and technical constraints, and power disparities. Ignoring these 'fractures' will only make AI governance harder.

Carabantes highlights Burrell's (2016) framework, illustrating that AI's opacity stems from three main layers. First, AI's complexity is intentionally enhanced - using methods that prioritise performance over transparency. Secondly, its workings are kept secret by businesses and governments to safeguard their algorithms, competitive edge, and reputation, leveraging legal protections. Third, the advanced nature of AI technology makes it inaccessible to the general population. This opacity primarily serves the interests of powerful entities that utilise AI for surveillance, influence, and financial gain. Furthermore, Big Tech employs AI to subtly guide behaviour towards certain outcomes, a strategy unlikely to diminish, suggesting an expected rise in surveillance and manipulation through AI.

## AI bias, normativity and discrimination

When AI models exhibit stereotypes, discrimination, and exclusion, the models are not inherently at fault, as they mirror the societal values prevalent in the context where they were developed. Rather than originating these values, the models are shaped by them and can further propagate and mask their societal roots. Machine learning models absorb and reflect the language and thought processes of their surrounding society, influencing it through their application. Bias, normativity, and discrimination will persist in models as long as these issues are present in the society from which the models learn.

Pop Stefanija (Chapter 49) shares her research into the data that a number of digital technology companies hold about her:

It demonstrated how little we know about who holds data about us, why they hold that data, and what they do with it. Crucially, it showed that, while there are ways to obtain information, to *make the invisible visible*, there are almost no ways of knowing how that data might and will affect us, or how to act agentially once information is obtained (p. 563).

Pop Stefania argues that it is crucial to remember that not all problems are suited for technological solutions, especially those stemming from deep-seated systemic power imbalances, which demand systemic responses rather than technical fixes. Before turning to AI for answers, we must first consider whether it is appropriate to involve AI in the decision-making process. Even if a computer suggests it is, we must retain the ability to challenge its authority and decisions. This includes questioning the rationale behind its outputs and, importantly, having the power to reject or alter its recommendations.

Brown (Chapter 50) intriguingly argues that the apparent absence of race in virtual assistants actually emphasises racial differences by ignoring non-white identities, which alienates users of colour while catering to the assumed preferences of white users. This approach to designing gendered and racialised tech products generates "the false possibility" of imagining an equitable, post-work world without first tackling the existing inequalities rooted in gendered and racialised divisions of labour (p. 581).

## Politics and activism in AI

The sixth and penultimate section of the Handbook deals with issues of politics, activism, and AI. It is the shortest section, with only five chapters. Eriksson's chapter (58) is one of several that deals with automation and the future of work. It focuses on the Swedish automation debate. Eriksson makes the important point that critical AI analysts need to urgently reflect on automation. This begins with debunking the myth that technological progress is an unstoppable force beyond our control and acknowledging the fundamental importance of politics.



Holzappel's chapter (60) is titled "Introducing political ecology of creative-AI". The author discusses various "creative support tools" and how they have been used in various art forms, such as music and literature (p. 691). He reminds us that a team of music historians, musicologists, composers, and computer scientists spent years preparing data and designing AI models to complete Beethoven's 10th Symphony.

## AI and automation in society

Technology is not necessarily good, nor bad; nor is it neutral. (Kranzberg, 1986, p. 545, cited on p. 726)

The final part of this Handbook encompasses a variety of chapters that apply a critical perspective to explore diverse realms and situations involving the development, usage, or deployment of AI. Lindberg's vision for these chapters is to collectively provide a thorough review of the present landscape of AI and automation within society and highlight ways it can be subject to critical examination.

Parviainen (Chapter 69) amuses us by reminding us of the disastrous launch of Mistubishi's Wakamaru domestic robot that did not lead to a single sale. Although there is a widespread belief that AI-powered robots will transform caregiving, their impact on nursing remains limited. Currently used technologies include monitoring devices, automated medication dispensers, robotic pets, mobile telepresence tools, and logistics support in hospitals. However, these devices mainly perform basic, routine interactions or simple repetitive tasks and are not equipped for complex multitasking or assisting with daily activities. I will mention in passing that Schiff and Rosenberg-Kima (Chapter 70) provide a useful overview of key milestones in the history of artificial intelligence in education (AIED) and AI.

Brevini (Chapter 75) adopts a comprehensive approach to AI's lifecycle that reveals its significant ecological costs. The process begins with the extraction of rare metals and minerals, critical for AI hardware, linked to technocolonialism and resulting in environmental and social harm, including damaged ecosystems and loss of biodiversity. (Technocolonialism is Madianou's (2019) term that refers to the convergence of digital developments with humanitarian structures, state power and market forces and the extent to which they reinvigorate and rework colonial genealogies.) AI production and operation further strain the environment through high energy use, emissions, and electronic waste, with cloud computing's carbon footprint notably exceeding that of the airline industry. Additionally, AI systems' water use for cooling data centres adds to their ecological impact. The cycle ends with the disposal of electronic waste, disproportionately affecting developing countries. This lifecycle analysis highlights the urgent need for sustainable and equitable AI practices, challenging both Big Tech and governments to address the environmental consequences of AI.

## Critique and conclusion

While some may find meta-critical thinking tiresome, critical theorists such as Marcuse and Angela Davis stress the importance of critical theory being critical of itself (Davis, 1989; Brookfield et al., 2024). In this spirit, some critical remarks can be made. The trouble with academic books is similar to that of journal articles: they lag behind what's happening. Consequently, generative AI does not have much presence in this Handbook; for instance, the index mentions ChatGPT only once.

Of the 127 authors, 123 are based in Western countries, and only two are based in South Africa and one in Japan and India each. Hence, the critique of Westocentrism or Eurocentrism could be considered. However, the difficulties in getting critical theory contributions on AI from non-Western countries must be considerable. In fact, Okolo's chapter (33) discusses the global inequality in AI and machine learning (ML) publications: Latin American, African and Southeast Asian countries are far behind the top publishing nations. It appears more problematic that there is no contribution from China, an AI superpower, and there are few mentions of China (the index claims that there is only one page (p. 217) that discusses China).

The Handbook's writing may, on occasion, be difficult to penetrate for readers who are not well-versed in critical theory, sociology and philosophy. In my view, that's okay, and there is no need to dumb down the Handbook's challenging topics. Einstein famously paraphrased Occam's Razor by saying that everything should be made as simple as possible, but not simpler. He emphasised the importance of simplicity in understanding complex ideas. Einstein's advice may not have been followed in the Handbook, and matters are occasionally portrayed in a more complex way than necessary. However, the cardinal sin of putting simple ideas into difficult language is rarely committed. More often than not, complex ideas are conveyed in difficult language. Helen Sword's advice, with which I struggle myself, is that stylish academic writers should gravitate toward "complex ideas communicated in clear, comprehensible language" (Sword, 2017, p. 152; see Green, 2009).

Finally, the book's price is the elephant in the room of this review. Given the Handbook's gargantuan size, it is, unsurprisingly, not a cheap book. However, it is laudable that the publisher, Edward Elgar, has not only made this Handbook available at the normal hardback price of £310 (the Edward Elgar membership price is £279), but e-book options for individuals start at a much more palatable £48. The book may be expensive, but there is enough material for weeks or even months of intensive reading in it. While it would be unethical to copy it, you could recommend it to your university or national library or share your copy with friends and colleagues.

In any event, I highly and unreservedly recommend this excellent Handbook. It emerges as an indispensable text for those immersed in digital sociology, science and technology studies and blends rich theoretical insights with empirical analyses. It is a vital resource for anyone keen to critically explore the complex relationship between AI and society.

Lindgren's Handbook serves as a critical guide for postgraduate students, scholars, practitioners, developers, and policymakers who want to acquire a deeper appreciation of the socio-political landscape of AI. In an era where the discourse around AI ethics often aligns with the interests of Big Tech, adopting a critical perspective toward AI becomes imperative. The Handbook's 76 chapters address a wide array of topics, from the political economy to socio-technological narratives and activism and present a critical exploration of AI's entanglement with social structures and power dynamics. It stands as a counter to the prevailing ideologies of technological optimism and solutionism, advocating for the development of technology that fosters, rather than hinders, societal well-being and communal harmony. Lindgren's Handbook not only aids in comprehending the current state of AI but also supports the critical and interdisciplinary endeavour to create technology that enhances collective welfare.

If you cannot access or afford the book or you are intimidated by its epic length, there are excellent, thinner and more affordable books that are critical on AI. Three of my favourites are Simon Lindgren's *Critical theory of AI* (2024), Kate Crawford's *Atlas of AI* (2021) and Stefan Popenici's *Artificial Intelligence and learning futures* (2023; see Rudolph, 2023).

### Additional references

Brookfield, S. D., Rudolph, J., & Tan, S. (2024). *Teaching well. Understanding key dynamics of learning-centered classrooms*. Routledge. <https://doi.org/10.4324/9781003447467>

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1). <https://doi.org/10.1177/2053951715622512>

Coeckelbergh, M. (2021). Time machines: Artificial intelligence, process, and narrative. *Philosophy & Technology*, 34(4), 1623-1638. <https://doi.org/10.1007/s13347-021-00479-y>

Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press. <https://doi.org/10.2307/j.ctv1ghv45t>

Forst, R. (2017). *Normativity and power: Analyzing social orders of justification*. Oxford University Press. <https://doi.org/10.1093/oso/9780198798873.001.0001>

Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. Basic Books.

Green, D. A. (2009). New academics' perceptions of the language of teaching and learning: Identifying and overcoming linguistic barriers. *International Journal for Academic Development*, 14(1), 33-45. <https://doi.org/10.1080/13601440802659254>

Hampton, G. J. (2015). *Imagining slaves and robots in literature, film, and popular culture: Reinventing yesterday's slave with tomorrow's robot*. Lexington Books.

Horkheimer, M. (1972). *Critical theory: Selected essays*. Continuum.

Kranzberg, M. (1986). Technology and history: "Kranzberg's laws". *Technology and Culture*, 27(3), 544-560. <https://doi.org/10.2307/3105385>

Lagerkvist, A. (2020). Digital limit situations: Anticipatory media beyond 'the new AI era'. *Journal of Digital Social Research (JDSR)*, 2(3), 16-41. <https://doi.org/10.33621/jdsr.v2i3.55>

Lindgren, S. (2024). *Critical theory of AI*. Polity.

MacCarthy, M. & Propp, K. (2021, May 4). *Machines learn that Brussels writes the rules. The EU's new AI regulation*. Brookings Institute. <https://www.brookings.edu/articles/machines-learn-that-brussels-writes-the-rules-the-eus-new-ai-regulation/>

Madianou, M. (2019). Technocolonialism: Digital innovation and data practices in the humanitarian response to refugee crises. *Social Media + Society*, 5(3), 2056305119863146

Popenici, S. (2023). *Artificial intelligence and learning futures. Critical narratives of technology and imagination in higher education*. Routledge. <https://doi.org/10.4324/9781003266563>

Rehak, R. (2021). The language labyrinth: Constructive critique on the terminology used in the AI discourse. In P. Verdegem (Ed.), *AI for everyone?* (pp. 87-102). University of Westminster Press. <https://doi.org/10.48550/arXiv.2307.10292>

Rheingold, H. (1994). *The virtual community: Homesteading on the electronic frontier*. Harper Perennial. <https://doi.org/10.7551/mitpress/7105.001.0001>

Rudolph, J. (2023). Book review. Popenici, Stefan (2023). *Artificial intelligence and learning futures. Critical narratives of technology and imagination in higher education*. Routledge. *Journal of Applied Learning and Teaching*, 6(2), 420-426, <https://doi.org/10.37074/jalt.2023.6.2.27>.

Rudolph, J., Ismail, M. F., & Popenici, S. (2024). Higher education's generative artificial intelligence paradox: The meaning of chatbot mania. *Journal of University Teaching and Learning Practice*, 21(6). Ahead of publication.

Searle, J. R. (1992). *The rediscovery of the mind*. MIT Press.

Sword, H. (2017). *Stylish academic writing*. Harvard University Press.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. <https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf>

Copyright: © 2024. Jürgen Rudolph. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Vol.7 No.1 (2024)

# Journal of Applied Learning & Teaching

ISSN : 2591-801X

Content Available at : <http://journals.sfu.ca/jalt/index.php/jalt/index>

Book review of Benedict du Boulay, Antonija Mitrovic, & Kalina Yacef (Eds., 2023). *Handbook of artificial intelligence in education*. Edward Elgar.

Jürgen Rudolph<sup>A</sup>

A

Director of Research & Learning Innovation, Kaplan Higher Education Academy

DOI: <https://doi.org/10.37074/jalt.2024.7.1.31>

## Introduction

Lately, Artificial Intelligence in Education (AIED) has gained significant attention due to what the editors perceive as “the impressive achievements of artificial intelligence in general, the much greater public awareness of the subject (both positive and negative) and the effects of COVID on educational practice” (p. 4). Even prior to the recent upsurge in the appeal of generative AI – a topic not addressed in the book due to its 2023 publication date – intelligent tutors engaged in listening and verbal interaction with learners, simulated student proficiency and understanding, examined learning theories, and facilitated classroom management, scalable learning, evaluation, and human-machine interaction.

The *Handbook of artificial intelligence in education* seeks to provide a comprehensive overview of the field by tracing its development from the 1970s to the present. Well-known authors in the field of AIED evaluate various AI techniques that are used to create systems that support learners, teachers, and administrators. Du Boulay et al.’s Handbook covers theories, foundational aspects, collaborative learning, games, and psychomotor learning. It concludes with a discussion of the wider context, commercialisation, social and political role, ethics, and future challenges and opportunities for AIED. It provides both broad and detailed accounts of the field. The target audience of the Handbook is researchers and (advanced) students in computer science, education, and AI.

The editors are well-known AI researchers and computer scientists. The lead editor, Benedict du Boulay, is one of the pioneers in AIED. He is a past president of the International Society for Artificial Intelligence in Education, an emeritus professor of Artificial Intelligence at the University of Sussex, and a visiting professor at University College London. Antonija Mitrovic is a Professor of Computer Science and Software Engineering at the University of Canterbury (New Zealand), and Kalina Yacef is a Professor at the School of Computer Science at the University of Sydney.

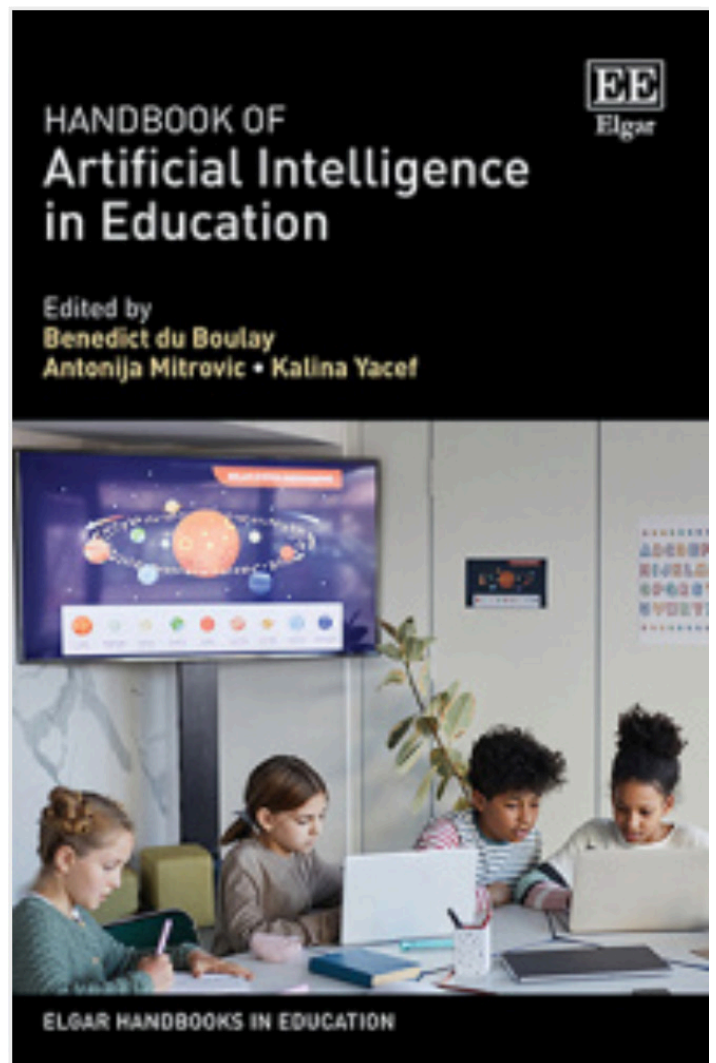


Figure 1: Book cover of du Boulay et al. (Eds., 2023).

The structure of du Boulay et al.’s exceptionally well-referenced Handbook represents their sense of the scope of AIED in six parts: (1) scene setting, (2) theories underpinning AIED, (3) the architecture and design of AIED systems, (4) analytics, (5) AIED systems in use and (6) the future. The first five parts of the 696-page Handbook consist of 26 chapters. The final part manifests the editors’ creative idea

of combining 11 brief visions of the future in the 27th multi-authored chapter. I undertake my best efforts to avoid writing a lengthy book review and hence limit myself to a cursory overview of the book before I provide some concluding remarks.

## Overview

In addition to the editors' helpful introduction, the initial section, "Scene setting", presents a comprehensive overview of AIED's young history. This journey began with the early adoption of computers in education, initially referred to as Computer-Assisted Learning (CAL) or Computer-Assisted Instruction (CAI). The field has grown exponentially in the past 30 years. Exploring AIED's history provides insights into how the discipline has evolved and the reasons behind the existence of certain research directions today. The saying, 'Those who are unaware of history may be destined to repeat it', highlights the critical role of historical awareness in preventing the recurrence of past errors and steering the course of future developments.

Part 2 discusses the foundational theories of AIED across four chapters, each focusing on different aspects: cognitive and knowledge theories; metacognition (an individual's awareness and understanding of their own thought processes); the realms of emotion and meta-emotion (how we think and feel about our emotional experiences); and the principle of scrutability. Ohlson's Chapter 3 examines the significance of psychological, instructional design, and AI theories in the development of AIED systems. Azevedo and Wiedbusch, in Chapter 4, detail six systems designed to enhance learners' metacognitive skills. Chapter 5, authored by Arroyo et al., examines the roles of emotions and motivation in learning. Kay et al., in Chapter 6, focus on the concept of "Scrutable AIED". Scrutability pertains to creating AIED systems that are accountable, transparent, explainable, and ethical. Scrutable AIED can be fully understood and evaluated through careful analysis.

Part 3 encompasses the foundational aspects of AIED system design, spread over six chapters. It starts with an exploration of three crucial elements for adaptive and personalised learning: domain modelling (structuring the subject content), student modelling (assessing learners' needs and knowledge), and instructional strategies (learning from problem-solving practice, learning from examples, exploratory learning, collaborative learning, game-based learning and learning by teaching). Following this, the discussion shifts to enhancing learner engagement through design and natural language communication, concluding with an analysis of the tools for AIED system development.

Part 4 encompasses five chapters dedicated to exploring analytics and data-driven approaches within AIED systems. Two of the chapters (Chapters 14 and 15) concentrate on aiding educators. Holstein and Olsen's chapter (14) explores how AI can collaborate with teachers in learning settings like classrooms, emphasising a partnership where AI assists through a human-in-the-loop and co-agency approach. Pozdniakov et al.'s chapter (15) examines various types of teacher-facing analytics and details their utility in sense-

making, decision-making, and pedagogical actions. Rosé et al.'s chapter (17) focuses on social analytics and foundational theories aimed at enhancing learning experiences in Massive Open Online Courses (MOOCs).

Part 5 examines the diverse learning modalities enabled by AIED and addresses the evaluative and assessment methodologies pertinent to these systems, their commercial trajectories, and the overarching ethical and bias-related discourse within the sector. The initial trio of chapters (18-20) scrutinises the support AIED provides for distinct learning types: psychomotor development, from assisting individuals with motor challenges to the training of surgeons; the dynamics of online collaborative learning; and the pedagogical potential of gamified learning environments. Subsequently, a pair of methodological chapters (21-22) is dedicated to assessment and evaluation. This is followed by another duo of chapters (23-24) that explore the commercialisation aspects of AIED technologies. The section culminates with two chapters (25-26) that critically engage with the social and ethical implications inherent to the deployment of AIED solutions.

To me, one of the Handbook's highlights is Luckin and Cukurova's chapter (24) on "Small-scale commercialisation: the golden triangle of AI EdTech". Together with two colleagues, I recently interviewed Prof Rose Luckin on her fascinating work, so I may be biased (Luckin et al., 2024). The Golden Triangle Framework outlines a collaborative model for AIED, linking educators, researchers, and EdTech developers to enhance teaching and learning. It underlines the critical role of educators in managing AI tools for personalised student learning experiences. The framework advocates for AI solutions that support teachers. It draws from multidisciplinary research for evidence-based, human-centred design. Luckin and Cukurova stress that ethical considerations, such as transparency and privacy, are essential for trust and learner autonomy. Their approach emphasises AI as a supportive co-pilot in education rather than being on auto-pilot.

Another important chapter (25) in Part 5 is Williamson et al.'s "Critical perspectives on AI in education: political economy, discrimination, commercialization, governance, and ethics". The chapter covers some of the ground of another recently published Handbook by Edward Elgar: Simon Lindgren's (Ed., 2023) *Handbook of critical studies of artificial intelligence*, which I strongly recommend reading (Rudolph, 2024). Williamson et al. provide a sharp focus on educational technology and their chapter ends with the following caution:

Depoliticized, ahistorical and asocial approaches are a continued problem in many studies of edtech, and enable those with limited expertise but significant (often commercial, economic or political) power to take centre stage in shaping and investing in educational futures (p. 566).

The multi-authored final section casts an eye toward the future. It highlights themes of access, equity, diverse learning modalities, enabling technologies, and the challenges of ubiquity and information overload. It aligns with the

United Nations Sustainable Development Goal #4, which is committed to achieving inclusive and equitable quality education and fostering lifelong learning for everyone. Despite this commitment, the reality remains stark, as highlighted by recent UNESCO statistics: 262 million children were out of school, two-thirds failed to learn basic literacy and numeracy skills after several years of schooling, and 750 million adults were illiterate. In WEIRD (Western, Educated, Industrialised, Rich, and Democratic) countries, the disparity is further exacerbated by the Matthew effect: those with access to advanced educational systems gain increasingly more knowledge, skills, and opportunities. Modern educational software, with its assumptions of readily available resources such as stable electricity, advanced computers, high-speed internet, and expert support, mirrors and magnifies these inequalities, predominantly benefiting those within WEIRD nations.

## Conclusion

Du Boulay et al.'s *Handbook of artificial intelligence in education* may well become a major new destination for the tech-savvy, techno-optimistic AI literati. Venturing into its pages as someone whose roots lie far from computer science, I confess the lexicon and prose style demanded a certain acclimatisation. Upon reflection, it became apparent that my intellectual palate found a more satisfying meal within the pages of another Edward Elgar scholarly feast: the *Handbook of critical studies on artificial intelligence* (Lindgren (Ed.), 2023; see Rudolph, 2024). The contributions on AIED – particularly those by Rahm (2023), Connolly (2023), and Schiff & Rosenberg-Kima (2023) – resonated with me more than the bulk of the current volume under scrutiny. However, I urge the esteemed readers of this review not to let my personal idiosyncratic academic inclinations – steeped as they are in the humanities, social sciences, and the realms of business and management – unduly colour their perception. In the final analysis, Du Boulay, Motrović, and Yacef's encyclopaedic Handbook may well be a timely and much-needed AIED magnum opus.

## Additional references

Connolly, R. (2023). From ethics to politics: Changing approaches to AI education. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 493-503). Edward Elgar.

Lindgren, S. (Ed.). (2023). *Handbook of critical studies of artificial intelligence*. Edward Elgar.

Luckin, R., Rudolph, J., Grünert, M., & Tan, S. (2024). Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin. *Journal of Applied Learning and Teaching*, 7(1). <https://doi.org/10.37074/jalt.2024.7.1.27>

Rahm, L. (2023). Educational imaginaries of AI. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 289-300). Edward Elgar Publishing Limited. <https://doi.org/10.4337/9781803928562.00031>

Rudolph, J. (2024). Book review of Lindgren, S. (Ed., 2023), *Handbook of critical studies of artificial intelligence*. *Journal of Applied Learning and Teaching*, 7(1). <https://doi.org/10.37074/jalt.2024.7.1.30>

Schiff, D. S., & Rosenberg-Kima, R. B. (2023). AI in education: Landscape, vision and critical ethical challenges in the 21st century. In S. Lindgren (Ed.), *Handbook of critical studies of artificial intelligence* (pp. 804-814). Edward Elgar.