

Impact and implications of AI methods and tools for the future of education

Edited by

Kingsley Okoye, Samira Hosseini, Kamal Kant Hiran
and Julius Nganji

Published in

Frontiers in Education
Frontiers in Computer Science



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-5013-7
DOI 10.3389/978-2-8325-5013-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Impact and implications of AI methods and tools for the future of education

Topic editors

Kingsley Okoye — Institute for the Future of Education, Monterrey Institute of Technology and Higher Education (ITESM), Mexico

Samira Hosseini — FEMSA Biotechnology Center, School of Engineering and Sciences, Monterrey Institute of Technology and Higher Education, Mexico

Kamal Kant Hiran — Symbiosis University of Applied Sciences, India

Julius Nganji — University of Toronto, Canada

Citation

Okoye, K., Hosseini, S., Hiran, K. K., Nganji, J., eds. (2024). *Impact and implications of AI methods and tools for the future of education*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-5013-7

Table of contents

- 05 **Editorial: Impact and implications of AI methods and tools for the future of education**
Kingsley Okoye, Julius T. Nganji, Kamal Kant Hiran and Samira Hosseini
- 09 **Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse?**
Kevin Fuchs
- 13 **A review of integrating AI-based chatbots into flipped learning: new possibilities and challenges**
Chung Kwan Lo and Khe Foon Hew
- 20 **Complex thinking through a Transition Design-guided Ideathon: testing an AI platform on the topic of sharing economy**
Jorge Sanabria-Z, Isolda Margarita Castillo-Martínez, Laura Icela González-Pérez and María Soledad Ramírez-Montoya
- 33 **Gains from the transition from face-to-face to digital/online modality to improve education in emergency**
Claudia Camacho-Zuñiga, Maritza Peña-Becerril, Maria de la O Cuevas-Cancino and Eréndira Gabriela Avilés-Rabanales
- 45 **How are AI assistants changing higher education?**
Eva-Maria Schön, Michael Neumann, Christina Hofmann-Stölting, Ricardo Baeza-Yates and Maria Rauschenberger
- 54 **Fostering digital transformation in education: technology enhanced learning from professors' experiences in emergency remote teaching**
Enrique Mondragon-Estrada, Ingrid Kirschning, Juan Arturo Nolasco-Flores and Claudia Camacho-Zuñiga
- 68 **The impact of ChatGPT on higher education**
Juan Dempere, Kennedy Modugu, Allam Hesham and Lakshmana Kumar Ramasamy
- 81 **AI-assisted analysis of content, structure, and sentiment in MOOC discussion forums**
Michael Yee, Anindya Roy, Meghan Perdue, Consuelo Cuevas, Keegan Quigley, Ana Bell, Ahaan Rungta and Shigeru Miyagawa
- 98 **Predictive analytics study to determine undergraduate students at risk of dropout**
Andres Gonzalez-Nucamendi, Julieta Noguez, Luis Neri, Víctor Robledo-Rella and Rosa María Guadalupe García-Castelán
- 112 **ChatGPT, a brand-new tool to strengthen timeless competencies**
Barthélémy Michalon and Claudia Camacho-Zuñiga

- 125 **ChatGPT in didactical tetrahedron, does it make an exception? A case study in mathematics teaching and learning**
Dadan Dasari, Agus Hendriyanto, Sani Sahara, Didi Suryadi, Lukman Hakim Muhaimin, Theodore Chao and Laila Fitriana
- 140 **Incorporating vertical collaboration to address sustainable development goals: the Monarch Route Project**
Maria de la O Cuevas-Cancino, Maritza Peña-Becerril, Enrique Mondragon-Estrada and Claudia Camacho-Zuñiga
- 151 **Evaluating effectiveness of immersive virtual reality in promoting students' learning and engagement: a case study of analytical biotechnology engineering course**
Angélica Lizeth Sánchez-López, Jesús Antonio Jáuregui-Jáuregui, Nancy Anabel García-Carrera and Yocanxóchitl Perfecto-Avalos
- 158 **From flexibility to adaptive learning: a pre-COVID-19 perspective on distance education in Latin America**
Alan Miralrio, Jessica Muñoz-Villota and Claudia Camacho-Zuñiga



OPEN ACCESS

EDITED AND REVIEWED BY
Clifford A. Shaffer,
Virginia Tech, United States

*CORRESPONDENCE
Kingsley Okoye
✉ kingsley.okoye@tec.mx

RECEIVED 17 May 2024
ACCEPTED 22 May 2024
PUBLISHED 31 May 2024

CITATION
Okoye K, Nganji JT, Hiran KK and Hosseini S
(2024) Editorial: Impact and implications of AI
methods and tools for the future of
education. *Front. Educ.* 9:1434052.
doi: 10.3389/feduc.2024.1434052

COPYRIGHT
© 2024 Okoye, Nganji, Hiran and Hosseini.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Editorial: Impact and implications of AI methods and tools for the future of education

Kingsley Okoye^{1,2*}, Julius T. Nganji³, Kamal Kant Hiran⁴ and Samira Hosseini^{1,5}

¹Writing Lab, Institute for Future of Education, Tecnologico de Monterrey, Monterrey, Nuevo Leon, Mexico, ²Department of Computer Science, School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey, Nuevo Leon, Mexico, ³Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada, ⁴Aalborg University, Copenhagen, Denmark, ⁵School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey, Nuevo Leon, Mexico

KEYWORDS

artificial intelligence, digitized-education, AI methods, intelligent systems, educational innovation, future of education, sustainable education, SDG4

Editorial on the Research Topic

Impact and implications of AI methods and tools for the future of education

Overview

The congruence of *technology* such as artificial intelligence (AI) and its use for *education* can help transform pedagogical practices and the future of education (Walsh, 2020; ICFE, 2021; UNESCO, 2023; Okoye et al., 2024a). Leading educational bodies like UNESCO (2021a,b, 2023) and the World Bank (2021) are advocating for research and development initiatives or endeavors focused on leveraging technology to reimagine education, and operationalize the use of AI tools and methods, commonly referred to as “digitized education.” This paradigm shift entails the formulation of methodological frameworks and guidelines to facilitate widespread adoption of AI methodologies in the educational settings (Nganji, 2018; UNESCO, 2019a,b; Arqawi et al., 2022; Okoye et al., 2024b). Also noteworthy is the fact that “digitized education” is now an inevitable and integral element to achieving the global sustainable development goals (SDGs) (UN, 2015; Pedró et al., 2019; UNESCO, 2019a; Okoye et al., 2021, 2022, 2024a; UIS, 2023) particularly the SDG4 that promotes quality of education (Global Goals, 2023).

This Research Topic is based on the premise that the different stakeholders, researchers, and AI practitioners, can commit to the creation, rethinking, and re-imagining of the AI methods and technologies that could result in effective adoption and implementation of TEL-based (technology-enhanced) education, science, and innovations. This collection of research aimed to collect state-of-the-art studies that looks into (i) why digitized-education may or may not be effectively implemented in different settings, particularly in education (ii) intelligent tools, models, and manifestos that support instructional AI design for scale or learning sustainability (iii) strategies, policies, and frameworks that supports AI-based interventions to facilitate the teacher-students learning processes and outcome (iv) digitally engaging the educational ecosystem, and (v) data-driven methods or analytical approaches for human learning development and life-long learning (UNESCO, 2019a; World Bank, 2021; UIS, 2023; Okoye et al., 2024a).

The topics in this collection encompass studies employing diverse methodologies, including case studies, exploration of emerging technologies like ChatGPT, and comprehensive reviews aimed at understanding the impact, benefits, barrier and bottleneck associated with digital technologies or AI tools in education. The articles collected for this Research Topic focus on providing the readers with wide knowledge of the “Use, impact and implications of AI methods and tools for the future of education” ranging from the ChatGPT and AI chatbots, predictive learning analytics, to transition from face-to-face to digital/online modalities of learning, distance learning pre- and -post COVID-19 pandemic, intelligent tutoring systems, virtual reality, augmented reality, flipped classrooms, and blended learning, challenges and opportunities with AI for digitized-education, e.g., students’ learning and engagement, ethics and policies, and digitized-education and its transformative implications for achieving the global sustainable development goals (SDGs), e.g., quality of education, inclusive education, and topic of sharing economy.

Spanning Original Research, Review and Opinion Articles, Methods, Hypothesis and Theory, and Curriculum, Instruction, and Pedagogy Articles; this Research Topic brings together leading scholars and researchers from diverse geographical regions and disciplinary backgrounds. By presenting the latest advancements and historical insights into the intersection of AI and Education, this Research Topic serves as a valuable resource for the global academic community and readership.

Articles in this Research Topic

The editors received a diverse range of submissions, including original research and review articles, focusing primarily on the “Use, impact and implications of AI methods and tools for the future of education.” In total, 23 manuscripts were submitted for consideration, each undergoing thorough evaluation by at least two to three independent experts in the field. Following a rigorous review process and assessment of relevance and scientific merit of the received manuscripts; 14 articles were chosen for inclusion in this Research Topic. These articles were authored by 52 esteemed experts from around the world, all renowned in their respective fields. Below is a concise overview of the accepted papers in this Research Topic:

In the paper by [Dempere et al.](#), the authors explored the effects of Artificial Intelligence (AI) chatbots, with a particular focus on OpenAI’s ChatGPT, on Higher Education Institutions (HEIs). The systematic review study revealed diverse perspectives on ChatGPT’s potential in education, and identified notable benefits to include research support, automated grading, and enhanced human-computer interaction. Conversely, risks such as privacy breaches, misuse, bias, misinformation, decreased human interaction, and accessibility issues were also identified in the review study.

In the paper by [Schön et al.](#), the authors discussed the impact of AI assistants in the context of higher education, outlined possible changes to the context, and present recommendations for adapting to change. The study reviewed related works in the topic and developed a conceptual structure that visualizes the role of AI assistants in higher education. They argued that AI assistants will change the context of higher education in a disruptive manner,

and note that the tipping point for this transformation has already been reached.

In the paper by [Dasari et al.](#), the authors adopted the didactical tetrahedron model, which addresses the issue of integrating technology into the didactical triangle, by examining the role of ChatGPT within educational settings. They found that students who rely solely on ChatGPT for learning resulted in lower performance in comparison to those receiving instruction from teachers, either alone or supported by ChatGPT. The study notes that the integration of ChatGPT into educational frameworks remains questionable within a didactic context, particularly due to its limitations in fostering deep information comprehension, stimulating critical thinking, and providing human-like guidance. It advocates for a balanced approach, suggesting that ChatGPT can augment the learning process effectively when used in conjunction with guidance.

In the paper by [Yee et al.](#) the authors propose an AI-assisted labeling process that uses advanced natural language processing techniques to train machine learning models capable of labeling a large dataset while minimizing human annotation effort. They fine-tune pretrained transformer-based deep learning models on category, structure, and emotion classification tasks. The study found significant differences in how learners of different age groups, gender, and course ask for help, provide help, and make posts with emotional (positive or negative) sentiment.

In the paper by [Lo and Hew](#), the authors investigate the emerging concept of integrating AI-based chatbots into flipped learning (chatbot-supported flipped learning) and its potential to enhance students’ learning experience. The findings of the review study suggest that this emerging instructional approach could result in benefits such as increased student interaction with learning content, improved class preparation, and data-driven teaching and learning. The review offers insights into future research and development in the topic area, and highlights the potential challenges in the topic to include limited technical functionality, lack of authenticity, and insufficient student motivation.

In the paper by [Gonzalez-Nucamendi et al.](#), the authors conducted quantitative research using machine learning tools to detect undergraduate students at risk of dropping out, and the factors associated with this phenomenon. In their approach, Clustering algorithms and Classification methods were tested to determine the predictive power of the different analyzed variables regarding the dropout/retention rate using data ($n = 14,495$) about the undergraduate students. The study found that the main variables associated with student dropouts were their academic performance during the early weeks of the first semester, their average grade in the previous academic levels, the previous mathematics score, and the entrance exam score.

In the paper by [Mondragon-Estrada et al.](#), the authors investigated professors’ perception of the key competences, and the best educational strategies and technological tools to guide digital transformation (DT) in education, according to their experience in emergency remote teaching (ERT). The outcome of the study showed social intelligence as an underlying competence for teaching performance, and that implementing information and communication technologies resulted in maintaining students’ interest and building trust in the online environment. It suggests

learning-by-doing, flipped learning, problem-based learning, game-based learning, and holistic education as some pedagogical methodologies that were successfully applied in ERT and could be implemented for DT.

In the paper by [Michalon and Camacho-Zuñiga](#), the authors investigated how ChatGPT was integrated in an undergraduate course for an International Relations program in a private Mexican university. The students were evaluated on their ability to explain to ChatGPT several discipline-specific methods and to make the AI implement these methods step by step. The outcome of the study which prove useful toward curriculum and educational policy development, shows that the students not only learned how to use the new AI tool and deepen their understanding of prospective methods, but also strengthened three soft or transversal competencies: communication, critical thinking, and logical and methodical reasoning.

In the paper by [Miralrio et al.](#), the authors uncovered the historical development of distance education (DE) in Latin America (LATAM) by drawing connections with five generations or milestones of the evolution of ICT in the region: Correspondence, Audiovisual, ICTs-based, Web-based, and Interactive. The review study of well-known International databases from end of nineteenth century to the COVID-19, confirms the critical role of Higher Education Institutions (HEIs) in LATAM in supporting ICT integration, and demonstrating a synergistic interplay with DE. The research highlights the value of DE as a cost-effective solution to traditional education, addressing social barriers and broadening educational access in LATAM, especially during emergencies.

In the paper by [Sánchez-López et al.](#), the authors conducted an immersive virtual reality (IVR) intervention in an analytical biotechnology course to evaluate its effectiveness in promoting student learning and academic engagement. The findings of their approach which included theoretical framework, IR spectra exercises, an individual quiz, and competition games; highlighted the effectiveness of IVR in enhancing learning outcomes, curiosity, motivation, and engagement among students. It also provides empirical evidence of IVR's positive impact on student learning and engagement in the context of biotechnology engineering.

In the paper by [Camacho-Zuñiga et al.](#), the authors examined the challenges associated with transitioning from face-to-face to online/digital lessons, according to perceptions of professors ($n = 105$) in a Mexican private university, few months after mandatory online/digital teaching was implemented. By analyzing data collected through a survey using statistical and data mining techniques, the authors identified factors that should be considered when designing future education in emergency situations. The study found that while most professors positively evaluated their performance in emergency remote teaching (ERT), they still preferred face-to-face teaching due to the lack of personal interaction and distractions. Best practices for online and face-to-face modalities to improve future education in emergencies were also discussed in the paper.

In the paper by [Sanabria-Z et al.](#), the authors apply a proof-of-concept study of an AI-based platform aimed to integrate a sequence of activities into the design of an online

platform to assess the development of complex thinking competency in higher education students. In their study, Transition Design method was used to explore the complex thinking meta-competency and its sub-competencies of critical, innovative, scientific, and systemic thinking in a sharing economy challenge for the students. The study found the two tested hypothesis: (a) the functioning of the constructed learning situation and (b) the model of the learning situation, to be true.

In the paper by [Cuevas-Cancino et al.](#), the authors adopted a research-based learning methodology to implement a horizontal (intra-class), vertical (inter-class), and interdisciplinary (multiple courses) student collaboration done by analyzing survey data on the perceived learning outcomes (87% confidence level) and letter for future classes through sharing their experience on The Monarch Route Project (MRP) on Sustainable Development Engineering (SDE) program. The study analysis which includes a mixed methodology (i.e., quantitative and qualitative, including text mining) evidenced that the participants strengthened both, transverse competences (collaboration, citizenship, communication, and critical thinking) and disciplinary competences (subject matter). It highlights the importance of integrating the UN's Sustainable Development Goals (SDGs) and the 2030 Agenda into the HEI's curricula.

In the paper by [Fuchs](#), the author theoretically discusses a range of challenges and opportunities with the generative AI tools such as Chat GPT for higher education, as well as the underlying implications that may expose gaps in the literature, stimulate research ideas, and, finally, advance the discussion about NLP in higher education. The article points to the potential benefits of using NLP models for personalized learning and on-demand support, such as providing customized learning plans, generating feedback and support, and offering resources to students whenever and wherever they need them. In addition, it highlights and discusses the challenges that NLP models may bring to include the potential loss of human interaction, bias, and ethical implications.

Final thoughts

The guest editors anticipate that this Research Topic, along with the accepted papers in the Research Topic, will not only benefit readers and researchers in AI and Education but also the broader scientific community interested in exploring the "Impact and implications of AI methods and tools for the future of education." This Research Topic serves as a valuable reference for educators, AI professionals, and PhD students, while also offering methodological guidance for educational policymakers and AI developers. The editors would like to thank the contributing authors for their insightful and impactful research. We would also like to thank the Editorial Office, Editor-in-Chief of the Digital Education Section, Editorial Manager and Assistance team, and the Expert Reviewers of the Frontiers in Education, for their immense and professional support toward the successful completion of this Research Topic "Impact and implications of the AI methods and tools for the future of education."

Author contributions

KO: Conceptualization, Project administration, Resources, Supervision, Validation, Writing—original draft. JN: Conceptualization, Project administration, Resources, Supervision, Validation, Writing—original draft. KH: Project administration, Resources, Supervision, Validation, Writing—review & editing. SH: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing—review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

References

- Arqawi, S. M., Zitawi, E. A., Rabaya, A. H., Abunasser, B. S., and Abu-Naser, S. S. (2022). Predicting university student retention using artificial intelligence. *Int. J. Adv. Comput. Sci. Appl.* 13, 315–324. doi: 10.14569/IJACSA.2022.0130937
- Global Goals (2023). *Goal 4: Quality Education - The Global Goals*. Available online at: <https://www.globalgoals.org/goals/4-quality-education/> (accessed March 17, 2022).
- ICFE (2021). “Reimagining our futures together: a new social contract for education - UNESCO Report by International Commission on the Futures of Education (ICFE),” in *Reimagining our Futures Together: A New Social Contract for Education*. Available online at: <https://unesdoc.unesco.org/ark:/48223/pf0000379707.locale=en> (accessed June 27, 2022).
- Nganji, J. T. (2018). Towards learner-constructed e-learning environments for effective personal learning experiences. *Behav. Inf. Technol.* 29:470673. doi: 10.1080/0144929X.2018.1470673
- Okoye, K., Daruich, S. D. N., Castaño, R., Enríquez de la, J. F., Escamilla, J., and Hosseini, S. (2024a). Analyzing the impact of digitized-education toward the future of education: a comparative study based on students' evaluation of teaching data. *Stu. Educ. Eval.* 82:101359. doi: 10.1016/j.stueduc.2024.101359
- Okoye, K., Hussein, H., Arrona-Palacios, A., Quintero, N., Luis, -P., Ortega, O. P., et al. (2022). Impact of digital technologies upon teaching and learning in higher education in Latin America: an outlook on the reach, barriers, and bottlenecks. *Educ. Inf. Technol.* 2022, 1–70. doi: 10.1007/s10639-022-11214-1
- Okoye, K., Nganji, J. T., Escamilla, J., and Hosseini, S. (2024b). Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education. *Comput. Educ. Artif. Int.* 12 :100205. doi: 10.1016/j.caeai.2024.100205
- Okoye, K., Rodriguez-Tort, J. A., Escamilla, J., and Hosseini, S. (2021). Technology-mediated teaching and learning process: a conceptual study of educators' response amidst the COVID-19 pandemic. *Educ. Inf. Technol.* 26, 7225–7257. doi: 10.1007/s10639-021-10527-x
- Pedró, F., Subosa, M., Rivas, A., and Valverde, P. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. Education Sector – UNESCO – ED-2019/WS/8. p. 46. Available online at: <https://unesdoc.unesco.org/ark:/48223/pf0000366994> (accessed May 26, 2022).
- UIS (2023). *Data for Education - A Guide for Policymakers to Leverage Education Data - UNESCO*. Available online at: <https://ces.uis.unesco.org/wp-content/uploads/sites/23/2024/01/Data-for-Education-final.pdf> (accessed January 23, 2024).
- UN (2015). *United Nations Sustainable Development – 17 Goals to Transform Our World*. Available online at: <https://www.un.org/sustainabledevelopment/> (accessed June 28, 2020).
- UNESCO (2019a). *Beijing Consensus on Artificial Intelligence and Education - UNESCO Digital Library*. Available online at: <https://unesdoc.unesco.org/ark:/48223/pf0000368303> (accessed March 11, 2022).
- UNESCO (2019b). *How can Artificial Intelligence Enhance Education? | UNESCO*. Available online at: <https://www.unesco.org/en/articles/how-can-artificial-intelligence-enhance-education> (accessed January 24, 2023).
- UNESCO (2021a). *Education and Technological Transformations for Human-Centered Recovery: The Global Education Coalition in Action - ED/GEC/2021/04*. Available online at: <https://unesdoc.unesco.org/ark:/48223/pf0000379133> (accessed October 28, 2021).
- UNESCO (2021b). *Education Research and Foresight*. Available online at: <https://en.unesco.org/themes/education/research-foresight> (accessed May 3, 2022).
- UNESCO (2023). *Artificial Intelligence and the Futures of Learning | UNESCO*. Available online at: <https://www.unesco.org/en/digital-education/ai-future-learning>
- Walsh, P. (2020). *Innovative Technology is the Future of Education*. Available online at: <https://www.forbes.com/sites/forbestechcouncil/2020/07/23/innovative-technology-is-the-future-of-education/?sh=16c19bfa70e3> (accessed July 7, 2020).
- World Bank (2021). *Digital Technologies in Education* [Image] | *Education and Technology Overview: Development News, Research, Data*. Available online at: <https://www.worldbank.org/en/topic/edutech#2> (accessed December 3, 2021).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



OPEN ACCESS

EDITED BY

Kamal Kant Hiran,
Symbiosis University of Applied Sciences, India

REVIEWED BY

Lina Kaminskiene,
Vytautas Magnus University, Lithuania

*CORRESPONDENCE

Kevin Fuchs
✉ kevin.f@phuket.psu.ac.th

RECEIVED 15 February 2023

ACCEPTED 11 April 2023

PUBLISHED 17 May 2023

CITATION

Fuchs K (2023) Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse? *Front. Educ.* 8:1166682. doi: 10.3389/feduc.2023.1166682

COPYRIGHT

© 2023 Fuchs. 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.

Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse?

Kevin Fuchs*

Faculty of Hospitality and Tourism, Prince of Songkla University, Phuket, Thailand

KEYWORDS

higher education, natural language processing, Chat GPT, technology, pedagogy

1. Introduction

The world has changed a lot in the past few decades, and it continues to change. Chat GPT has created tremendous speculation among stakeholders in academia, not the least of whom are researchers and teaching staff (Biswas, 2023). Chat GPT is a Natural Language Processing (NLP) model developed by OpenAI that uses a large dataset to generate text responses to student queries, feedback, and prompts (Gilson et al., 2023). It can simulate conversations with students to provide feedback, answer questions, and provide support (OpenAI, 2023). It has the potential to aid students in staying engaged with the course material and feeling more connected to their learning experience. However, the rapid implementation of these NLP models, like Chat GPT by OpenAI or Bard by Google, also poses several challenges. In this article, I will discuss a range of challenges and opportunities for higher education, as well as conclude with implications that (hopefully) expose gaps in the literature, stimulate research ideas, and, finally, advance the discussion about NLP in higher education.

2. Discussion

2.1. The emergence of NLP in academia

Natural Language Processing (NLP) models have been in development since the 1950s (Jones, 1994) but it was not until the past decade that they gained significant attention and advancement, particularly with the development of deep learning techniques and large datasets (Kang et al., 2020). NLP models are rapidly becoming relevant to higher education, as they have the potential to transform teaching and learning by enabling personalized learning, on-demand support, and other innovative approaches (Odden et al., 2021). In higher education, NLP models have significant relevance for supporting student learning in multiple ways. These models can be employed to analyze and process vast amounts of textual data, such as academic papers, textbooks, and other course materials, to provide students with personalized recommendations for further study based on their learning requirements and preferences. In addition, NLP models can be used to develop chatbots and virtual assistants that offer on-demand support and guidance to students, enabling them to access help and information as and when they need it.

Chat GPT by OpenAI and Bard (Google's response to Chat GPT) are examples of NLP models that have the potential to transform higher education. These generative language models, i.e., Chat GPT and Google Bard, can generate human-like responses to open-ended prompts, such as questions, statements, or prompts related to academic material. The recent release and increasing popularity (in early 2023) of Chat GPT and Google Bard made its use particularly relevant for supporting student learning in a range of contexts, such as language learning, writing, research, and general academic inquiry. Therefore, the use of NLP models in higher education expands beyond the aforementioned examples, with new applications being developed to aid students in their academic pursuits.

2.2. Opportunities for higher education

Personalized learning is an approach to education that aims to tailor instruction to the unique needs, interests, and abilities of individual learners. NLP models can facilitate personalized learning by analyzing students' language patterns, feedback, and performance to create customized learning plans that include content, activities, and assessments tailored to the individual student's needs. Personalized learning can be particularly effective in improving student outcomes. Research has shown that personalized learning can improve academic achievement, engagement, and self-efficacy (Wu, 2017). When students are provided with content relevant to their interests and abilities, they are more likely to engage with the material and develop a deeper understanding of the subject matter. NLP models can provide students with personalized learning experiences by generating content tailored specifically to their individual learning needs.

For example, when a student submits a response to a question, the model can analyze the response and provide feedback customized to the student's understanding of the material. This feedback can help the student identify areas where they might need additional support or where they have demonstrated mastery of the material. Furthermore, the processing models can generate customized learning plans for individual students based on their performance and feedback. These plans may include additional practice activities, assessments, or reading materials designed to support the student's learning goals. By providing students with these customized learning plans, these models have the potential to help students develop self-directed learning skills and take ownership of their learning process.

Moreover, on-demand support is a crucial aspect of effective learning, particularly for students who are working independently or in online learning environments. The NLP models can provide on-demand support by offering real-time assistance to students struggling with a particular concept or problem. The benefits of on-demand support are numerous. It can help students overcome learning obstacles and enhance their understanding of the material. In addition, on-demand support can help build students' confidence and sense of self-efficacy by providing them with the resources and assistance they need to succeed. These models can offer on-demand support by generating responses to student queries and feedback in real time. When a student submits a

question or response, the model can analyze the input and generate a response tailored to the student's needs.

This can be particularly helpful for students working independently or in online learning environments where they might not have immediate access to a teacher or tutor. Furthermore, chatbots can offer support to students at any time and from any location. Students can access the system from their mobile devices, laptops, or desktop computers, enabling them to receive assistance whenever they need it. This flexibility can help accommodate students' busy schedules and provide them with the support they need to succeed. Additionally, NLP models can provide students with on-demand support in a variety of formats, including text-based chat, audio, or video. This can cater to students' individual learning preferences and provide them with the type of support that is most effective for them.

2.3. Challenges for higher education

Although there is a wide range of opportunities for NLP models, like Chat GPT and Google Bard, there are also several challenges (or ethical concerns) that should be addressed. The first challenge is the issue of accuracy. The accuracy of the system depends heavily on the quality, diversity, and complexity of the training data, as well as the quality of the input data provided by students. In previous research, Fuchs (2022) alluded to the importance of competence development in higher education and discussed the need for students to acquire higher-order thinking skills (e.g., critical thinking or problem-solving). The system might struggle to understand the nuances and complexities of human language, leading to misunderstandings and incorrect responses. Moreover, a potential source of inaccuracies is related to the quality and diversity of the training data used to develop the NLP model.

If the training data is not adequately diverse or is of low quality, the system might learn incorrect or incomplete patterns, leading to inaccurate responses. The accuracy of NP models might be impacted by the complexity of the input data, particularly when it comes to idiomatic expressions or other forms of linguistic subtlety. Additionally, the model's accuracy might be impacted by the quality of the input data provided by students. If students do not provide clear, concise, and relevant input, the system might struggle to generate an accurate response. This is particularly challenging in cases in which students are not sure what information they need or cannot articulate their queries in a way that the system easily understands.

Another significant challenge that students might face when using NLP models in higher education is the potential risk of over-reliance on technology, which could undermine the development of important critical thinking skills (while critical thinking has been singled out as an exemplary skill, the list of skills is countless and multiple higher-order thinking skills could be further discussed in the context of chatbots). While these models can offer valuable support and personalized learning experiences, students must be careful to not rely too heavily on the system at the expense of developing their own analytical and critical thinking skills. Over-reliance on systems such as Chat GPT and Google Bard could lead to students becoming passive learners who simply accept the

responses generated by the system without questioning or critically evaluating the accuracy or relevance of the information provided. This could lead to a failure to develop important critical thinking skills, such as the ability to evaluate the quality and reliability of sources, make informed judgments, and generate creative and original ideas.

Moreover, over-reliance could reinforce existing biases and perpetuate inequalities in education. For example, if the system is trained on biased or incomplete data, it might generate responses reflecting these biases, thereby leading to a reinforcement of existing inequalities and a failure to challenge and disrupt discriminatory practices in higher education. To address these challenges, institutions must provide clear guidance to students on how to use NLP models as a tool to support their learning rather than as a replacement for critical thinking and independent learning. Institutions must also ensure that students are provided with opportunities to engage in active learning experiences that encourage critical thinking, problem-solving, and independent inquiry.

Another important challenge that should be mentioned is the linguistic aspect of NLP, like Chat GPT and Google Bard. Emerging evidence in the body of knowledge indicates that chatbots have linguistic limitations (Wilkenfeld et al., 2022). For example, a study by Coniam (2014) suggested that chatbots are generally able to provide grammatically acceptable answers. However, at the moment, Chat GPT lacks linguistic diversity and pragmatic versatility (Chaves and Gerosa, 2022). Still, Wilkenfeld et al. (2022) suggested that in some instances, chatbots can gradually converge with people's linguistic styles. While the development of artificial intelligence and natural language processing models like Chat GPT is just the beginning (Molnár and Szűts, 2018), it is not far-fetched to hypothesize that over time the linguistic accuracy of NLP models will improve and more closely mimic the writing style of humans (including expressive writing styles as similarly alluded to by Park et al., 2021).

3. Conclusion

In this article, I discussed the challenges and opportunities regarding natural language processing (NLP) models like Chat GPT and Google Bard and how they will transform teaching and learning in higher education. The article highlights the potential benefits of using NLP models for personalized learning and on-demand support, such as providing customized learning plans, generating feedback and support, and offering resources to students

whenever and wherever they need them. However, the article also acknowledges the challenges that NLP models may bring, including the potential loss of human interaction, bias, and ethical implications. To address the highlighted challenges, universities should ensure that NLP models are used as a supplement to, and not as a replacement for, human interaction. Institutions should also develop guidelines and ethical frameworks for the use of NLP models, ensuring that student privacy is protected and that bias is minimized.

Additionally, universities should involve students in the development and implementation of NLP models to address their unique needs and preferences. Finally, universities should invest in training their faculty to use and adapt to the technology, as well as provide resources and support for students to use the models effectively. In summary, universities should consider the opportunities and challenges of using NLP models in higher education while ensuring that they are used ethically and with a focus on enhancing student learning rather than replacing human interaction. Overall, NLP models are a powerful tool for improving the quality of education by providing students with personalized learning experiences and automating administrative tasks, while institutions must tackle the previously mentioned challenges to safeguard high-quality education for their students.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Biswas, S. (2023). ChatGPT and the future of medical writing. *Radiology* 307, 223312. doi: 10.1148/radiol.223312
- Chaves, A. P., and Gerosa, M. A. (2022). "The impact of chatbot linguistic register on user perceptions: a replication study," in *Chatbot Research and Design: 5th International Workshop, CONVERSATIONS 2021, Virtual Event* (Cham: Springer International Publishing) 143–159. doi: 10.1007/978-3-030-94890-0_9
- Coniam, D. (2014). The linguistic accuracy of chatbots: usability from an ESL perspective. *Text Talk*. 34, 545–567. doi: 10.1515/text-2014-0018
- Fuchs, K. (2022). The importance of competency development in higher education: Letting go of rote learning. *Front. Educ.* 7, 1004876. doi: 10.3389/feduc.2022.1004876
- Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., et al. (2023). How does chatgpt perform on the united states medical licensing examination? The implications of large language models for medical education and knowledge assessment. *JMIR Med. Educ.* 9, e45312. doi: 10.2196/45312
- Jones, K. S. (1994). "Natural language processing: a historical review," in *Current issues in Computational Linguistics* 3–16. doi: 10.1007/978-0-585-35958-8_1

Kang, Y., Cai, Z., Tan, C. W., Huang, Q., and Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *J. Manag. Analyt.* 7, 139–172. doi: 10.1080/23270012.2020.1756939

Molnár, G., and Szűts, Z. (2018). “The role of chatbots in formal education,” in *2018 IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY)* (IEEE) 000197–000202. doi: 10.1109/SISY.2018.8524609

Odden, T. O. B., Marin, A., and Rudolph, J. L. (2021). How has Science Education changed over the last 100 years? *An analysis using natural language processing. Sci. Educ.* 105, 653–680. doi: 10.1002/sce.21623

OpenAI (2023). *OpenAI Official Website. Introducing ChatGPT - Learn more.* Available online at: <https://openai.com/blog/chatgpt/> (accessed February 10, 2023).

Park, S., Thieme, A., Han, J., Lee, S., Rhee, W., and Suh, B. (2021). “‘I wrote as if I were telling a story to someone I knew’: Designing Chatbot Interactions for Expressive Writing in Mental Health,” in: *Designing Interactive Systems Conference 2021* 926–941. doi: 10.1145/3461778.3462143

Wilkenfeld, J. N., Yan, B., Huang, J., Luo, G., and Algas, K. (2022). “‘AI love you’: Linguistic convergence in human-chatbot relationship development,” in *Academy of Management Proceedings* (Briarcliff Manor, NY: Academy of Management) 17063. doi: 10.5465/AMBPP.2022.17063abstract

Wu, J. Y. (2017). The indirect relationship of media multitasking self-efficacy on learning performance within the personal learning environment: Implications from the mechanism of perceived attention problems and self-regulation strategies. *Comput. Educ.* 106, 56–72. doi: 10.1016/j.compedu.2016.10.010



OPEN ACCESS

EDITED BY

Kingsley Okoye,
Tecnológico de Monterrey, Mexico

REVIEWED BY

Yinghui Shi,
Central China Normal University, China
Ayesha Kanwal,
University of Glasgow, United Kingdom

*CORRESPONDENCE

Chung Kwan Lo
✉ chungkwanlo@eduhk.hk

RECEIVED 28 February 2023

ACCEPTED 04 May 2023

PUBLISHED 22 May 2023

CITATION

Lo CK and Hew KF (2023) A review
of integrating AI-based chatbots into flipped
learning: new possibilities and challenges.
Front. Educ. 8:1175715.
doi: 10.3389/feduc.2023.1175715

COPYRIGHT

© 2023 Lo and Hew. 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.

A review of integrating AI-based chatbots into flipped learning: new possibilities and challenges

Chung Kwan Lo^{1*} and Khe Foon Hew²

¹Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong, Hong Kong SAR, China, ²Faculty of Education, The University of Hong Kong, Hong Kong, Hong Kong SAR, China

This mini review examines the emerging concept of integrating AI-based chatbots into flipped learning and its potential to enhance students' learning experience. We investigate the design and practice of chatbot-supported flipped learning, as well as the benefits and challenges associated with this approach. Despite an extensive database search, only 10 empirical articles were found, indicating that this cutting-edge research topic requires further exploration. The findings of this review suggest that this emerging instructional approach could result in benefits such as increased student interaction with learning content, improved class preparation, and data-driven teaching and learning. However, potential challenges included limited technical functionality, lacking authenticity, and insufficient student motivation. The review offers insights into future research and development to advance the knowledge and practice of integrating AI chatbots into flipped learning.

KEYWORDS

flipped learning, flipped classroom, chatbot, technology-enhanced learning, literature review

Introduction

Artificial intelligence (AI) conversation chatbots have gained significant attention worldwide, especially after the release of ChatGPT by OpenAI¹ on November 30, 2022. Although ChatGPT has created an impact on different disciplines, it is widely reported that it relies on biased data and may provide incorrect or fake information (Lo, 2023). Therefore, there is still a need to build chatbots for specific purposes, such as guiding student learning in a course. An AI-based chatbot is a computer program designed to simulate human conversation through natural language processing to understand and respond to user queries in a human-like manner. These chatbots are thus valuable tools in various industries, from customer service (Nicolescu and Tudorache, 2022) to healthcare (Xu et al., 2021), by improving workflow efficiencies, reducing costs, and enhancing user experience.

In the education sector, Wollny et al. (2021) have classified AI chatbots into three major types: assisting chatbots (e.g., course assistance), mentoring chatbots (e.g., increasing self-awareness), and learning chatbots which are specifically trained to support students' learning of course materials. For example, Lee and Yeo (2022) developed an AI-based chatbot to act as a virtual student who had difficulty learning mathematics. This chatbot

¹ <https://openai.com>

was designed to enhance pre-service teachers' questioning skills through a 24/7 simulation of student-teacher dialogue.

There is a great potential to enhance flipped learning with AI chatbots (Diwanji et al., 2018). Flipped learning is an instructional approach that reverses the traditional order of class activities (Bishop and Verleger, 2013). In a typical flipped lesson, students engage with learning materials (e.g., instructional videos and readings) before class and then come to class prepared with basic knowledge to participate in more advanced and interactive learning activities (Hew et al., 2021a; Lo and Hew, 2022). Although recent meta-analyses have suggested that flipped learning could increase student achievement across subject disciplines (Cheng et al., 2019; van Alten et al., 2019; Hew et al., 2021a), its practices are not without challenges. As Akçayır and Akçayır (2018) found, two common complaints of flipped learning are that students lack guidelines at home, and that they are unable to get help during pre-class learning, which impedes their engagement in subsequent in-class activities. Ideally, AI chatbots can be used to provide them with 24/7 assistance and personalized support (Diwanji et al., 2018).

However, integrating AI chatbots into flipped learning is a new and emerging concept. While there have been studies on the use of AI chatbots in education (see Wollny et al., 2021 for a review) and flipped learning (see Hew et al., 2021a for a review) separately, existing reviews have found a scarcity of studies on their integration. In a recent systematic review by Wollny et al. (2021) on chatbots in education, only one study of flipped learning (i.e., Huang et al., 2019) was retrieved as of December 23, 2020. This indicates that it is still a relatively new research area worthy of a follow-up literature review, as the new development of AI chatbot technology may have led to more effective ways of using them in flipped learning and new challenges that need to be addressed. Hence, this mini review aims to understand recent practices for designing and integrating AI chatbots in flipped learning, their benefits, and challenges. Our findings can serve as a foundation for further research and development in this area. The following research questions (RQ1 to RQ3) are posed to guide this review.

- RQ1: How are AI chatbots designed and used in flipped learning?
- RQ2: What are the benefits of integrating AI chatbots into flipped learning?
- RQ3: What are the challenges of integrating AI chatbots into flipped learning?

Methods

Search strategies

We followed the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) statement (Moher et al., 2009) when selecting relevant articles. Eleven electronic databases were used, including (1) Academic Search Ultimate, (2) ACM Digital Library, (3) APA PsycInfo, (4) British Education Index, (5) CINAHL, (6) Education Research Complete, (7) ERIC, (8) IEEE Xplore, (9) PubMed, (10) Scopus, and (11) Web of Science. The

search string was designed based on recent literature reviews of AI chatbots (Wollny et al., 2021; Nicolescu and Tudorache, 2022) and flipped learning (Hew et al., 2021a; Lo and Hew, 2022), and it was formulated as follows: ("conversational AI" OR "dialogue system*" OR "dialog system*" OR "chatbot*" OR "conversational agent*") AND (flip* OR invert*) AND (class* OR learn* OR instruction* OR course*). The Boolean operators and asterisks were used to increase the flexibility of the search string in retrieving relevant articles of AI chatbots and flipped learning. A snowballing procedure was further executed using Google Scholar. Specifically, we tracked the research items which cited the included articles.

Inclusion and exclusion criteria

We conducted our final search on 30 April 2023 without specifying a period of publication. Therefore, relevant articles published before the search date were identified and screened. To answer our research questions, only empirical studies were included. However, no constraints were imposed on the types of empirical data used (e.g., surveys and interviews). To ensure consistency, the interventions had to satisfy the definition of flipped learning (i.e., students first engage with learning materials before class and then participate in in-class activities; Cheng et al., 2019; Lo and Hew, 2022) and use at least one chatbot in their practice. Considering the integration of AI chatbots and flipped learning is an emerging research area, we included conference papers and book chapters in addition to journal articles. No constraints were imposed on the location of interventions, research contexts, and language of instruction. However, the articles must be written in English. Table 1 summarizes the inclusion and exclusion criteria for study selection.

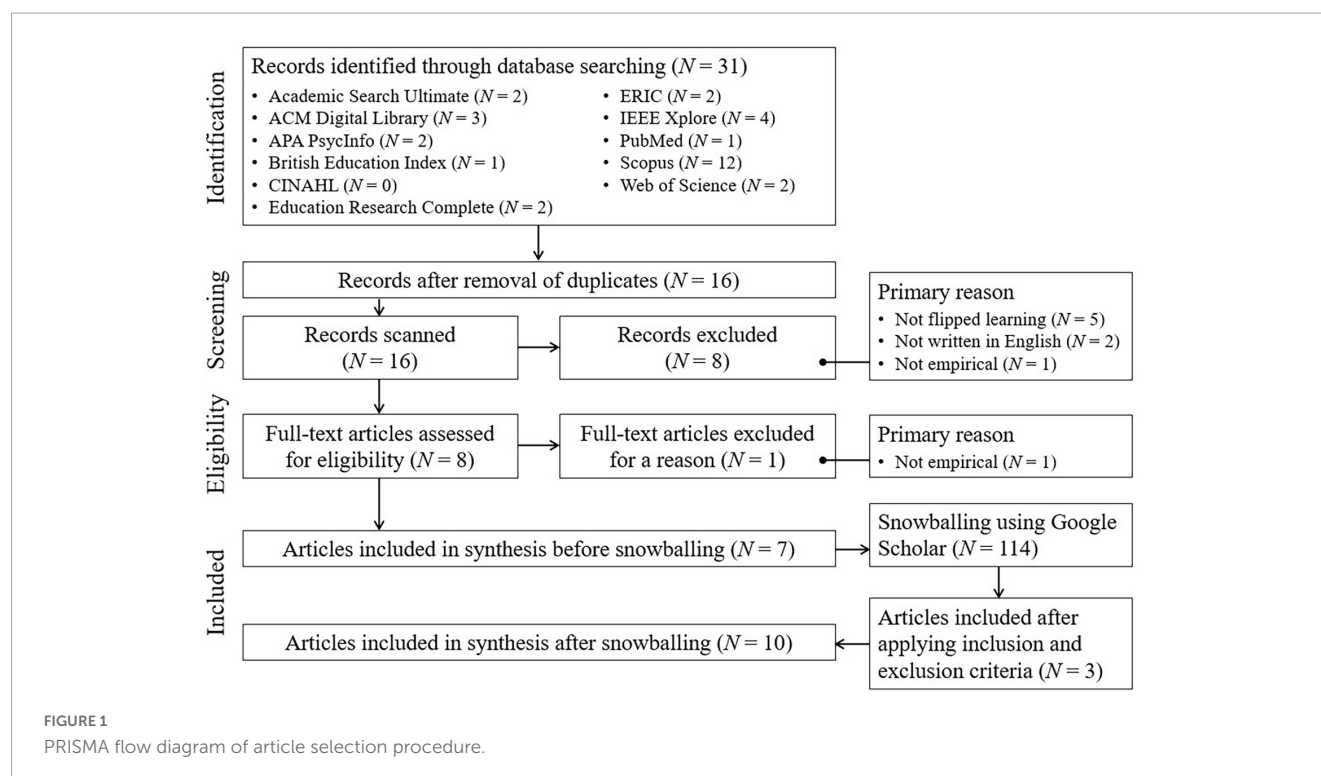
Data extraction and analysis

We extracted data from each article, including the authors(s), the year of publication, the country of implementation, the subject area, and the research participants. We further extracted information on how AI chatbots were designed and used in flipped learning (RQ1), how the integration benefited teachers and students (RQ2), and the challenges to the integration (RQ3). The data were analyzed through content analysis (Creswell, 2012), and emerging themes were identified through open coding. Similar codes were grouped and organized into themes. Exemplary quotations were identified to illustrate each constructed theme. The articles were double-coded to establish coding reliability, and disagreements were resolved through discussion among the authors.

Findings and discussion

Study selection and characteristics of the included articles

A total of 31 records were obtained through a database search as of 30 April 2023. After removing duplicates, there were 16



unique records. However, some were excluded because they were not related to flipped learning ($N = 5$), not written in English ($N = 2$), nor empirical studies ($N = 1$). One study was excluded after full-text assessment for eligibility due to a lack of empirical data. Nevertheless, it was used for background reference. Hence, seven articles were yielded. We then tracked the research items which cited the included articles using Google Scholar. After executing this snowballing procedure, three additional articles were included. Overall, the final selection yielded 10 articles for synthesis. Notably, Du et al. (2021) stated that their study was a follow-up study of Hew et al. (2021b). Therefore, a total of nine unique studies were analyzed. Figure 1 outlines the process of article selection.

The background of the reviewed studies is summarized in the Appendix. A majority of studies were conducted in Asia, such as Hong Kong (Gonda and Chu, 2019; Huang et al., 2019; Hew et al., 2021b, 2023; Li et al., 2021), Japan (Ito et al., 2021), and Taiwan (Lin and Mubarak, 2021). With the exception of the study by Timpe-Laughlin et al. (2022), all other studies involved students in higher education. However, the reviewed studies were conducted in various subject areas. Only the studies by Lin and Mubarak (2021), Timpe-Laughlin et al. (2022), and Hew et al. (2023) were within the same subject area (i.e., English language education).

RQ1: how are AI chatbots designed and used in flipped learning?

In the reviewed studies, only Lin and Mubarak (2021) adopted a ready-made chatbot application, called Replika, which allowed students to interact with and practice their English speaking. Researchers in other studies built their own chatbots for flipped learning. For example, Varnavsky (2022) used Python

programming to develop his chatbot and incorporated it with Telegram (an instant messaging application). Li et al. (2021) detailed their system parameters and variables considered (e.g., students' fraction of correctly answered pre-class quizzes and the time needed) when developing their chatbot. The chatbot was then implemented on Facebook Messenger and Telegram. Besides, Gonda and Chu (2019) and Hew et al. (2021b, 2023) built their chatbots using an existing platform, namely Google Dialogflow, whereas Huang et al. (2019) used another platform called IBM Watson Assistant. They had to customize the chatbots by setting intents (i.e., users' possible questions), entities (i.e., keywords which help the chatbots recognize users' words), and dialogue relevant to their courses.

Consistent with Wollny et al. (2021), the chatbots involved in the reviewed studies could be classified into three categories as

TABLE 1 Inclusion and exclusion criteria for study selection.

Criterion	Inclusion	Exclusion
The definition of flipped learning	Satisfy the definition of flipped learning (i.e., students first engage with learning materials before class and then participate in in-class activities)	Do not satisfy the definition of flipped learning
The use of chatbots	Use at least one chatbot into flipped learning	Do not use any chatbots in flipped learning
Time period	On or before 30 April 2023	Articles published after 30 April 2023
Study type	Empirical studies	Non-empirical studies
Document type	Journal articles, conference papers, and book chapters	Articles from other sources (e.g., media reports)
Language	English	Non-English

follows (the total number is greater than nine because the chatbots in some studies served multiple purposes).

- Learning chatbot ($N = 9$): The chatbots in all reviewed studies were used to facilitate students' out-of-class learning. They could provide real-time feedback on student performance (Gonda and Chu, 2019; Huang et al., 2019; Hew et al., 2021b, 2023; Li et al., 2021), answer students' questions (Gonda and Chu, 2019; Huang et al., 2019; Ito et al., 2021; Varnavsky, 2022), and allow students to practice their English speaking (Lin and Mubarak, 2021; Timpe-Laughlin et al., 2022).
- Assisting chatbot ($N = 2$): Varnavsky's (2022) chatbot could serve as a course assistant, such as "Providing material or links to material to be studied in preparation for the current workshop" (p. 291). In Li et al. (2021), "Students using Facebook Messenger chatbot in 2019 received their messages (i.e., quizzes and prompts) automatically (just like the Telegram user)" (p. 124).
- Mentoring chatbot ($N = 2$): We found two mentoring chatbots, namely "Self-Regulated Learning Chatbot" (Hew et al., 2021b, p. 169) and "goal-setting chatbot" (Hew et al., 2023, p. 40). Both of these helped students set their personal learning goals and then provided recommendations for them to achieve their goals.

RQ2: what are the benefits of integrating AI chatbots into flipped learning?

Several reviewed studies provided evidence that integrating chatbots into flipped learning was useful (Hew et al., 2021b, 2023) and enhanced student learning (Li et al., 2021). The following five specific benefits were identified.

- Immediate feedback ($N = 4$): Chatbots could provide students with timely feedback to support their learning and achievement of learning goals (Gonda and Chu, 2019; Huang et al., 2019; Hew et al., 2021b). In the words of one student, the chatbot provided prompt feedback which "solved students' problems in time" (Hew et al., 2021b, p. 173). Ito et al. (2021) further noted that students could ask their chatbot during class and even at midnight when they did their homework.
- Increased students' interaction with learning content ($N = 4$): Chatbot-integrated learning led to an increase in students' interaction with course materials (Huang et al., 2019; Hew et al., 2021b, 2023). In their English-speaking course, Lin and Mubarak (2021) further used mind maps to guide students' interactions with their chatbot, resulting in a significant increase in student-chatbot interaction as evidenced by the greater number of words produced.
- Improved students' class preparation ($N = 3$): The use of chatbots sending reminders to students led to a significant decrease in the percentage of students who did not read pre-class materials (Varnavsky, 2022). Li et al. (2021) found that around 90% of their students completed their pre-class quizzes within 3 days with messages pushed in their chatbots. Gonda and Chu (2019) further noted that their chatbot

could encourage their student to think and ask questions outside the classroom.

- Increased confidence ($N = 2$): This benefit was identified in the studies of English-speaking courses. Speaking with chatbots could boost students' confidence, as perceived by the teacher participants of Timpe-Laughlin et al. (2022). Similarly, Lin and Mubarak (2021) reported that their students felt more confident in their conversations.
- Data-driven teaching and learning ($N = 1$): The chatbot of Li et al. (2021) collected and analyzed students' performance in pre-class and in-class quizzes. Based on the analytics, the chatbot could provide pre-class quizzes that matched their level of proficiency and notify the teacher to modify the level of difficulty for learning tasks.

RQ3: what are the challenges of integrating AI chatbots into flipped learning?

While integrating chatbots into flipped learning can provide the aforementioned benefits, several challenges were identified in the reviewed studies, as follows.

- Limited technical functionality ($N = 4$): Hew et al. (2021b, 2023) pre-defined several options in their chatbot to assist students in labeling their input. However, it appeared that the variety of options was inadequate to fully meet the needs of their students. Huang et al. (2019) also noted that their chatbots could not solve unstructured problems requested by their students. In the study of Timpe-Laughlin et al. (2022), some teacher participants experienced difficulty with the chatbot's speech recognition and response accuracy.
- Lacking authenticity ($N = 2$): Timpe-Laughlin et al. (2022) reported concerns about the authenticity of learning tasks when using chatbots. The limited function of their chatbot might have restricted task complexity. Echoed with Timpe-Laughlin et al. (2022), one student of Huang et al. (2019) expressed that "I only want to ask very conceptual questions (what is...) from the chatbots, and as for some more complex questions (why..., how...), I prefer to ask a human friend" (p. 817).
- Lacking students' learning motivation ($N = 2$): Varnavsky (2022) reported that not all students used the chatbot throughout the course, and he attributed it to students' interest and motivation in the subject discipline. Similarly, Ito et al. (2021) found it challenging to motivate students during online learning, even with the presence of a chatbot.

Limitations and recommendations for future research and development

Two main limitations in this review must be acknowledged. First, although multiple databases were searched with a flexible

search string, only 10 relevant empirical studies could be found. While this review could provide valuable insights into the use of AI chatbots in flipped learning, the small number of reviewed studies limited the universality and generalizability of our findings. Moreover, because the reviewed studies were conducted in diverse subject areas, it is difficult to draw broad conclusions in this research topic. Hence, we recommend future studies be conducted in a range of subject areas (e.g., mathematics and health professions education) and contexts (e.g., K-12 schools) to advance our knowledge of the potential benefits and challenges of integrating AI chatbots into flipped learning.

Second, the findings on the effectiveness and usefulness of integrating AI chatbots into flipped learning were largely based on self-reported data (e.g., surveys and interviews). Only one study used achievement tests to assess students' learning outcomes (Lin and Mubarak, 2021). Therefore, we currently know little about the effect of using chatbots in flipped learning as compared with some other learning environments (e.g., flipped learning without using a chatbot). Further experimental research is required to examine the effectiveness of chatbot-supported flipped learning.

To address the challenges identified in this review, we provide an agenda for future research and development. First, it is necessary to develop a comprehensive database of Q&A data to enhance the chatbot's technical functionality and provide a wider range of responses to student queries (Ito et al., 2021). Involving frontline teachers in the chatbot and learning task design process can help ensure the authenticity of learning tasks and enhance the chatbot's technical functionality (Timpe-Laughlin et al., 2022). They can provide subject-specific insights into the types of questions and tasks that are relevant to their subject areas, which can become the basis for training data. In addition, to improve student motivation and engagement, allocating scores on learning tasks might be able to serve as an incentive to encourage class preparation (Li et al., 2021; Lo and Hew, 2022). However, to enhance intrinsic motivation, Diwanji et al. (2018) recommended the use of Self-determination Theory (Ryan and Deci, 2000) in chatbot and learning task designs. To enhance student engagement with chatbots, one student of Hew et al. (2021b) suggested embedding emojis in the chatbot responses in their chatbot. This suggestion echoed with the attempt of Lee and Yeo (2022) who made the responses of their chatbot more human-like by adding emojis and emotional reactions.

References

- Akçayır, G., and Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Comput. Educ.* 126, 334–345. doi: 10.1016/j.compedu.2018.07.021
- Bishop, J. L., and Verleger, M. A. (2013). "The flipped classroom: A survey of the research," in *Proceedings of 2013 ASEE annual conference & exposition*, (Atlanta, GA: ASEE), 6219. doi: 10.18260/1-2-22585
- Cheng, L., Ritzhaupt, A. D., and Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis. *Educ. Tech. Res. Dev.* 67, 793–824. doi: 10.1007/s11423-018-9633-7
- Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research.*, 4th Edn. Boston, MA: Pearson Education, Inc.

Conclusion

This mini review on integrating AI chatbots into flipped learning is significant in advancing our understanding of the state-of-the-art in this topic and its potential to enhance the learning experience for students. Our review identified the benefits (e.g., improved class preparation and data-driven teaching and learning) and potential challenges (e.g., limited technical functionality and lacking authenticity), which can inform educators and instructional designers in their implementation of chatbot-supported flipped learning. Most importantly, we provided recommendations for future research and development to address the challenges identified in this review. Nevertheless, more studies are required to investigate the effectiveness of chatbot-supported flipped learning in various educational contexts.

Author contributions

Both authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Funding

This research was funded by the Department of Mathematics and Information Technology, the Education University of Hong Kong, reference code MIT/DRG01/22-23.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Diwanji, P., Hinkelmann, K., and Witschel, H. F. (2018). "Enhance classroom preparation for flipped classroom using AI and analytics," in *Proceedings of the 20th international conference on enterprise information systems – volume 1: ICEIS*, (Funchal: Science and Technology Publications, Lda), 477–483. doi: 10.5220/0006807604770483
- Du, J., Huang, W., and Hew, K. F. (2021). "Supporting students goal setting process using chatbot implementation in a fully online course," in *Proceedings of 2021 IEEE international conference on engineering, technology & education*, (Wuhan: IEEE), 35–41. doi: 10.1109/TALE52509.2021.9678564
- Gonda, D. E., and Chu, B. (2019). "Chatbot as a learning resource? Creating conversational bots as a supplement for teaching assistant training course," in *Proceedings of 2019 IEEE international conference on engineering, technology and education*, (Yogyakarta: IEEE), 1–5. doi: 10.1109/TALE48000.2019.9225974
- Hew, K. F., Bai, S., Dawson, P., and Lo, C. K. (2021a). Meta-analyses of flipped classroom studies: A review of methodology. *Educ. Res. Rev.* 33:100393. doi: 10.1016/j.edurev.2021.100393
- Hew, K. F., Huang, W., Du, J., and Jia, C. (2021b). "Using chatbots in flipped learning online sessions: Perceived usefulness and ease of use," in *Blended learning: re-thinking and re-defining the learning process*, eds R. Li, S. K. S. Cheung, C. Iwasaki, L. F. Kwok, and M. Kageo (Cham: Springer), 164–175. doi: 10.1007/978-3-030-80504-3_14
- Hew, K. F., Huang, W., Du, J., and Jia, C. (2023). Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions. *J. Comput. High. Educ.* 35, 40–68. doi: 10.1007/s12528-022-09338-x
- Huang, W., Hew, K. F., and Gonda, D. E. (2019). Designing and evaluating three chatbot-enhanced activities for a flipped graduate course. *Int. J. Mech. Eng. Rob. Res.* 8, 813–818. doi: 10.18178/ijmerr.8.5.813-818
- Ito, T., Tanaka, M. S., Shin, M., and Miyazaki, K. (2021). "The online PBL (project-based learning) education system using AI (artificial intelligence)," in *DS 110: Proceedings of the 23rd international conference on engineering and product design education*, (Herning: Design Society), doi: 10.35199/EPDE.2021.19
- Lee, D., and Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Comput. Educ.* 191:104646. doi: 10.1016/j.compedu.2022.104646
- Li, J., Ling, L., and Tan, C. W. (2021). "Blending peer instruction with just-in-time teaching: Jointly optimal task scheduling with feedback for classroom flipping," in *Proceedings of the Eighth ACM conference on learning @ scale*, (Germany: ACM), 117–126. doi: 10.1145/3430895.3460134
- Lin, C. J., and Mubarak, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educ. Technol. Soc.* 24, 16–35. doi: 10.30191/ETS.202110_24(4).0002
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Educ. Sci.* 13:410. doi: 10.3390/educsci13040410
- Lo, C. K., and Hew, K. F. (2022). Design principles for fully online flipped learning in health professions education: A systematic review of research during the COVID-19 pandemic. *BMC Med. Educ.* 22:720. doi: 10.1186/s12909-022-03782-0
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and The Prisma Group. (2009). Reprint-preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Phys. Ther.* 89, 873–880. doi: 10.1093/ptj/89.9.873
- Nicolescu, L., and Tudorache, M. T. (2022). Human-computer interaction in customer service: The experience with AI chatbots—A systematic literature review. *Electronics* 11:1579. doi: 10.3390/electronics11101579
- Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/0003-066X.55.1.68
- Timpe-Laughlin, V., Sydorenko, T., and Daurio, P. (2022). Using spoken dialogue technology for L2 speaking practice: What do teachers think? *Comput. Assist. Lang. Learn.* 35, 1194–1217. doi: 10.1080/09588221.2020.1774904
- van Alten, D. C., Phielix, C., Janssen, J., and Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educ. Res. Rev.* 28:100281. doi: 10.1016/j.edurev.2019.05.003
- Varnavsky, A. N. (2022). "Chatbot to increase the effectiveness of the «flipped classroom» technology," in *Proceedings of 2022 2nd international conference on technology enhanced learning in higher education*, (Lipetsk: IEEE), 289–293. doi: 10.1109/TELE55498.2022.9801001
- Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., and Drachslar, H. (2021). Are we there yet? A systematic literature review on chatbots in education. *Front. Artif. Intell.* 4:654924. doi: 10.3389/frai.2021.654924
- Xu, L., Sanders, L., Li, K., and Chow, J. C. (2021). Chatbot for health care and oncology applications using artificial intelligence and machine learning: Systematic review. *JMIR Cancer* 7:e27850. doi: 10.2196/27850

Appendix

TABLE A1 Background of the reviewed studies.

Study (year of publication)	Location	Subject area	Participants (number)
Gonda and Chu, 2019	Hong Kong	Teaching assistant training	PG students ($N = 300$)
Du et al., 2021 ; Hew et al., 2021b	Hong Kong	Social media	PG students ($N = 15$ in course 1; $N = 29$ in course 2)
Hew et al., 2023	Hong Kong	Course 1: Engaging adult learners (course 1) Course 2: English listening (as foreign language)	Course 1: PG students ($N = 29$) Course 2: UG students ($N = 38$)
Huang et al., 2019	Hong Kong	Engaging adult learners	PG students ($N = 13$)
Ito et al., 2021	Japan	Project design	UG students ($N = 1,442$ in 2019; $N = 896$ in 2020)
Li et al., 2021	Hong Kong	Artificial intelligence	UG students ($N = 137$ in 2019; $N = 134$ in 2020)
Lin and Mubarak, 2021	Taiwan	English speaking (as foreign language)	UG students ($N = 22$ in C-AI group; $N = 28$ in MM-AI group)
Timpe-Laughlin et al., 2022	USA	English speaking (as foreign language)	Teacher participants ($N = 16$)
Varnavsky, 2022	Russia	Project documentation	Appeared to be UG students ($N = 80$)

PG, postgraduate; UG, undergraduate; C-AI, conventional AI chatbot; MM-AI, mind map-guided AI chatbot.



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Amro Al-Said Ahmad,
Keele University, United Kingdom
Christian Noumi,
University of Toronto, Canada

*CORRESPONDENCE

Isolda Margarita Castillo-Martínez
✉ isoldamcm@hotmail.com

RECEIVED 15 March 2023

ACCEPTED 26 April 2023

PUBLISHED 30 May 2023

CITATION

Sanabria-Z J, Castillo-Martínez IM,
González-Pérez LI and
Ramírez-Montoya MS (2023) Complex thinking
through a Transition Design-guided Ideathon:
testing an AI platform on the topic of sharing
economy.
Front. Educ. 8:1186731.
doi: 10.3389/feduc.2023.1186731

COPYRIGHT

© 2023 Sanabria-Z, Castillo-Martínez,
González-Pérez and Ramírez-Montoya. 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.

Complex thinking through a Transition Design-guided Ideathon: testing an AI platform on the topic of sharing economy

Jorge Sanabria-Z^{1,2}, Isolda Margarita Castillo-Martínez^{3*}, Laura Icela González-Pérez⁴ and María Soledad Ramírez-Montoya¹

¹Institute for the Future of Education, Tecnológico de Monterrey, Monterrey, Mexico, ²Department of Design, School of Architecture, Art and Design (EAAD), Tecnológico de Monterrey, Guadalajara, Mexico, ³School of Humanities and Education, Tecnológico de Monterrey, Monterrey, Mexico, ⁴Faculty of Psychology, Universidad Autónoma de Nuevo León, San Nicolás de los Garza, Mexico

Rationale: The development of the complex thinking meta-competency in the education of university students potentially promotes high capacities, where artificial intelligence (AI) might work as a supporting structure.

Objective: This proof-of-concept study of an AI-based platform aimed to integrate a sequence of activities into the design of an online platform to assess the development of complex thinking competency in higher education students.

Method: The Transition Design method was used within an Ideathon event supported by an AI platform to provide students with a sharing economy challenge. A total of 31 university students from five university campuses in Mexico synchronously participated. An instrument was applied in the pre-test and post-test phases to explore the complex thinking meta-competency and its sub-competencies of critical, innovative, scientific, and systemic thinking.

Results: Two hypotheses were tested, both of which turned out to be true: (a) the functioning of the constructed learning situation and (b) the model of the learning situation.

Conclusion: These findings may be of value to scientific and academic communities, as well as social and business organizations interested in developing high capacities of complex thinking and educational innovations using digital platforms.

KEYWORDS

complex thinking, artificial intelligence, Ideathon, Education 4.0, higher education

1. Introduction

Educational research uses methods that support the analysis of educational trends and priorities in the digital age to develop teaching and learning environments to properly train students to navigate digital transformation and address societal needs. Before the outbreak of COVID-19, Bonfield et al. (2020) provided different possible futures desirable in higher education, including smart campuses, digital assistants, lifelong learning, and online learning. Later, Yang et al. (2022) studied trends in global and digital education, revealing a rapid development phase (September 6, 2018–2022) where the research

hotspots of digital education primarily focused on interdisciplinary fields of practice and adaptive education research supported by Big Data. They predicted that human-computer interdisciplinary teaching models and smart education might become a future development trend of digital education. Other research throughout the pandemic emphasized ensuring continuous and universal education; each country needs to deploy a national (cluster) e-learning platform and ensure free access for all students (Ivanova et al., 2021) to increase the level of education of the population and guarantee continuity of knowledge throughout life (Krasovskiy et al., 2020). The contributions of knowledge generated by educational research in digital, remote, hybrid, and traditional formats should involve educational actors and tasks that include educational policies, changes in the structure of academic programs, new pedagogical approaches, updating of teachers, and the integration of interdisciplinary research groups to capitalize on the use of new technologies.

The global horizon of higher education and the paradigm shifts in educational trends require incorporating technologies that effectively solve problems through virtual and traditional learning interactions that positively impact citizens' skills. Artificial intelligence (AI) refers to the ability of a digital machine to perform tasks commonly associated with intelligent humans, and its associated technologies are divided into various fields, such as computer vision, speech, machine learning, Big Data, and natural language processing (Chiu, 2021). AI is incorporated into these research technologies to assist in predicting student performance and behavior, especially in platform online education. In particular, machine learning makes predictions using mathematical and statistical operations. In this case, it is convenient to analyze the data obtained from the education processes and evaluate the students' success and the factors affecting success (Akmeşe et al., 2021). For Yu (2021), online learning behavior refers to the learning-related behaviors in the network learning environment. Constructing a predictive model of academic performance in online education requires algorithms based on different languages supporting machine learning methods, regression, clustering, and preprocessing modules.

This paper aims to present the integrated flow of proof-of-concept activities for an AI platform that assesses the development of complex thinking meta-competency in higher education students. The event for running the test was a pedagogical intervention based on an Ideathon, assembled through a methodological sequence based on the Transition Design approach. Studies related to this research because of their approach incorporating AI to assess student behavior are those of Yu (2021), which shows an algorithm that uses a prediction method to help teachers and students conduct better teaching and learning activities. Another research that refers to this was conducted by Hu (2021), highlighting the use and optimization of machine learning algorithms. In turn, McGinnis et al. (2018) implemented the Scikit-learn (sklearn) toolkit based on Python, which contains various commonly used machine learning methods that facilitate analyzing a data set. Furthermore, Jarke and Macgilchrist, 2021 focused on the data dashboards of learning support systems based on machine learning (ML) and how these systems produce credible knowledge and compelling, persuasive, and convincing narratives as a pedagogical approach.

In this paper, the distinctive contribution is the methodology used, which involves Transition Design as a pedagogical approach to identify learning behaviors through an Ideathon-style event, with the primary objective of assessing participants' level of mastery of the

meta-competency of complex thinking when performing learning activities on an online AI platform. The results are presented for two hypotheses through statistical data revealing students' level of complex thinking at the end of the Transition Design activities.

2. Theoretical framework

2.1. Complex thinking in higher education

Higher education is moving toward rethinking the competencies needed for socio-economic, cultural, and environmental solutions that integrate emerging technologies. Changes in technology, social life and economics call for a change of traditional teaching and research methods (Bengu et al., 2020). Complex thinking is considered a meta-competency comprising four sub-competencies or types of thinking (critical, innovative, scientific, and systemic) (Ramírez-Montoya et al., 2022), which have been successfully applied and measured, for instance, in social entrepreneurial contexts (Vázquez-Parra et al., 2022). Critical thinking can be defined as an individual thought process that begins with the intent to solve a problem or answer a question by examining different options and choosing the most suitable and logical one (Alsaleh, 2020). Innovative thinking is the capacity for creativity, implemented with a high degree of success; four levels of innovation are delineated: incremental, modular, architectural, and radical (Passig and Cohen, 2014). Scientific thinking involves intentional information seeking, including asking questions, testing hypotheses, making observations, recognizing patterns, and making inferences (Kuhn, 2002). Systemic thinking uses methodological tools to manage emerging complexity in local and global contexts (Barile et al., 2018). Developing competencies and motivating learning involves tracking through the different stages of pedagogical approaches that integrate various technologies and consider the emotional support that should be provided to students during their training.

Fostering reasoning for complexity in higher education means enhancing skills to provide solutions to the challenges posed by new digital, social, environmental, and economic interactions. In higher education, complex thinking skills can develop students' competencies, human potential, and the capacity for innovation to solve new problems (Suárez-Brito et al., 2022). The digital transformation trends in teaching and learning are becoming increasingly oriented toward online learning (Marks and Al-Ali, 2022). By the end of 2021, Massive Open Course Online (MOOC) learning platforms enrollment exceeded 220 million students; 950 universities worldwide had announced or launched 19,400 courses (Shah, 2021). According to Cornejo-Velazquez et al. (2020), MOOC platforms offer value propositions to the universities and instructors, such as solid infrastructure in the cloud, marketing, advertising, and other administrative activities that allow for reducing operational costs of maintenance and updating. Although traditional higher education providers remain the dominant institutions, have the best reputation and are where most students aspire to go, MOOCs can work in a complementary way to strengthen the learning that higher education institutions seek to promote, as they are resources designed to respect the learning pace of students, as well as motivation through the inclusion of interactive elements.

Education worldwide is transforming and must consider economic and technological megatrends to connect with the skills and

competencies that students need. On the one hand, a world of possibilities opens up for us with AI. [Xia et al. \(2022\)](#) point to the potential of AI in several critical educational domains, such as (1) teaching, (2) learning, (3) evaluation, and (4) administration; but, on the other hand, [Munir and Awan \(2022\)](#) emphasize the importance of cultivating creative thinking, while considering emotions and intuition. This opens the door to new perspectives and avoids creating solutions based on scientific reasoning alone. One of the guidelines for developing a synergistic relationship between the AI functionalities and the sub-competencies of complex thinking is creating taxonomies of future scenarios through diverse, innovative teaching models where the learner's process is personalized and collaborative to contribute to the social construction of knowledge.

2.2. Creative design with AI in Education 4.0

AI in education (AIEd) gives rise to new opportunities by having structured and unstructured data, introducing the concept of Big Data, robotics, and artificial intelligence, to analyze and process for greater clarity of learning processes. The digital transformation of education allows systems to improve educational quality ([Chen, 2021](#)) and analyze the positive or negative effects of teaching by combining computer algorithms ([Hou, 2018](#)) and other technologies. [Liang and Wang \(2020\)](#) propose to improve the digital management of teaching through algorithms at various times with data analysis techniques. Also, teaching creativity contributes valuable solutions to any type of problem, including specifying the requirements of a new software or system ([Giunta et al., 2022](#); [Ramírez-Montoya et al., 2022](#)). Hybrid learning is gradually taking place in higher education, demanding new teaching and teacher interaction methods.

Various technologies with AIEd monitor and track student and teacher interactions and administrative tasks. Students and teachers accept AIEd because they can perceive that its compatibility can solve their learning needs ([Alhumaid et al., 2023](#)). AIEd has enormous potential to improve learning, teaching, assessment, and educational administration and provide anywhere, anytime machine-supported queries and immediate feedback ([Xia et al., 2022](#)). [Costa-Mendes et al. \(2021\)](#) argued that selecting student data for classical statistics may not fit AIEd predictive models to probe student performance. Predicting student performance with AIEd in contexts where creative design is taught essentially encompasses inquiring into students' development of reasoning for complexity at the neurological level.

Most e-learning platforms or learning management systems (LMS) offer functionalities to analyze the interactions of students and teachers to create statistics that give clues to the behavior of the actors. [Hamal et al. \(2022\)](#) emphasize that one of the most common applications of AI in education is intelligent tutoring systems (ITS), which can determine step-by-step an optimal path through support and learning activities; they are integrated with three models: (a) domain model represents the knowledge intended for students to acquire, (b) the pedagogical model represents knowledge of effective teaching and learning approaches that have been obtained from pedagogical experts and researchers in the learning sciences ([Siemens, 2012](#)), and (c) the learner model refers to the initial representation of the learner's state of knowledge. It is ideal to have these three models

to create AIEd data structures that provide more information on the aspects that need to be addressed to improve student performance.

2.3. Models that support AIEd

Taking into account the intelligent tutoring system created by [Hamal et al. \(2022\)](#), the three proposed models are described below: (a) domain model, (b) pedagogical model, and (c) learner model, which were considered for this research.

2.3.1. The domain model: sharing economy

With the emergence of Technology 4.0 enablers, new opportunities have arisen to exchange goods, services, and knowledge timely and collaboratively, without intermediaries. The phenomenon of enabling technologies for individuals or organizations to share goods or physical assets and reduce costs is called sharing economy ([World Economic Forum, 2015](#)). It aims to increase efficiency and optimize societal resources ([Muñoz and Cohen, 2017](#)), complementing the definition of [Wang and Ho \(2017\)](#) "... an emerging social and technological phenomenon based on developments in information and communications technology (ICT) that implies the collaborative consumption of physical, virtual, and intellectual goods." [Acquier et al. \(2017\)](#) posit the sharing economy on three fundamental pillars: (1) access economy, (2) platform economy, and (3) community-based economy. The most developed countries have invested in innovation, science, and technology, leading the growth of collaborative technology platforms that reflect new forms of sustainable consumption and have an impact far beyond their borders.

2.3.2. The pedagogical model: Transition Design

Transition Design is one of the pedagogies emerging to teach design with a focus on sustainability. For [Di Bella \(2022\)](#), Transition Design is "a new area of research, study, and practice, whose heuristic model is composed of (a) vision, (b) theories of change, (c) mindset and posture, and (d) new forms of design, which constitute the framework that defines four interrelated areas of knowledge, action, and self-reflection." Transition Design, as a pedagogical approach, could support and facilitate social transition processes by supporting, connecting, or developing interventions to change values, technologies, social practices, and infrastructures intentionally ([Irwin, 2015](#)). The need to create a future with a sustainable vision arises from the figure of the transition designer ([Irwin, 2012](#)), who must be an actor who provides solutions to the world engaged in complex systems that require a cultural transformation ([Di Bella, 2018](#)) considering economic megatrends and technologies. The pedagogical approach of Transition Design using technologies such as AI may provide clues as to what innovative solutions may be optimal for design education.

2.3.3. The learner model: Ideathon

Identifying the personal characteristics, attitudes, and behavior of students who use learning platforms offers knowledge about their behavior in different learning activities. Aligned with the United Nations' Sustainable Development Goals ([United Nations, 2015](#)), the Ideathon program aims at undergraduate students between 18 and 23 years old to foster potential change agents through the generation of innovative solutions to the challenges of our society. Ultimately, the Ideathon seeks to promote an ecosystem of high-impact

entrepreneurship in the early stages and foster a culture of innovation through access to knowledge, tools, expert mentoring, and talent linkage, in line with the global trends in innovation and technology (Haro, 2018). However, consolidating an online learning community requires considering the personalities of its members, the attitudes and values that underpin their practices, and their consumption of products and services (Bäro et al., 2022). Evaluating each student's learning based on their interactions and providing them with feedback and activities that allow them to reinforce specific skills is one of the functionalities expected to be solved through AI in learning platforms.

3. Hypotheses development

We hypothesized that since Ideathon-type events tend to trigger ideas and initiatives due to the collaborative energy and involvement they generate, this setting could provide the ideal environment for students between 18 and 23, who usually participate in open events. The focus on problem analysis and solution finding following the Transition Design methodology for a limited number of hours and in a competition-style environment (albeit without winners) led us to our first hypothesis:

H1. A Transition Design-driven Ideathon can be an engaging scenario to develop complex thinking in higher education students.

Secondly, we argued that student behaviors on the platform could be distributed into sections to be tracked by a computer system. Moreover, we envisioned that such behaviors could be recorded and accumulated by the system throughout the event and classified according to specific evaluation criteria, which led us to our second hypothesis.

H2. AI-provided digital platforms can measure the development of complex thinking traits in higher education students.

4. Methodology

Transition Design was the methodology employed in the educational scenario presented in this study (Irwin, 2015). According to Irwin et al. (2022), this approach allows an understanding of wicked problems typified by a diversity of stakeholders and concerns at different spatiotemporal layers, which, to be understood, requires multi-disciplinary and longitudinal interventions. Transition Design is emerging in the global north along with resilient Thinking and Policy Design, all pursuing plurality and synthesis of knowledge in systemic transformation processes (Juri et al., 2021). Examples of its application include the search for solutions to wicked problems faced by the Museum of Environmental Sciences in the framework of the "HUMETAV" project (Sanabria-Zepeda and Santana-Castellón, 2022); the generation of experiential futures in the field of fashion (Coward and Maione, 2022); and building collaborative media for the transformation of designers' mindsets (Bosch-Gómez et al., 2022). Irwin et al. (2022) describe it as a sequence of six steps that begins by (1) identifying the wicked problems and (2) their stakeholders, (3) mapping their historical origin, (4) creating their desirable long-term vision, (5) designing a pathway from the present to the future, and (6)

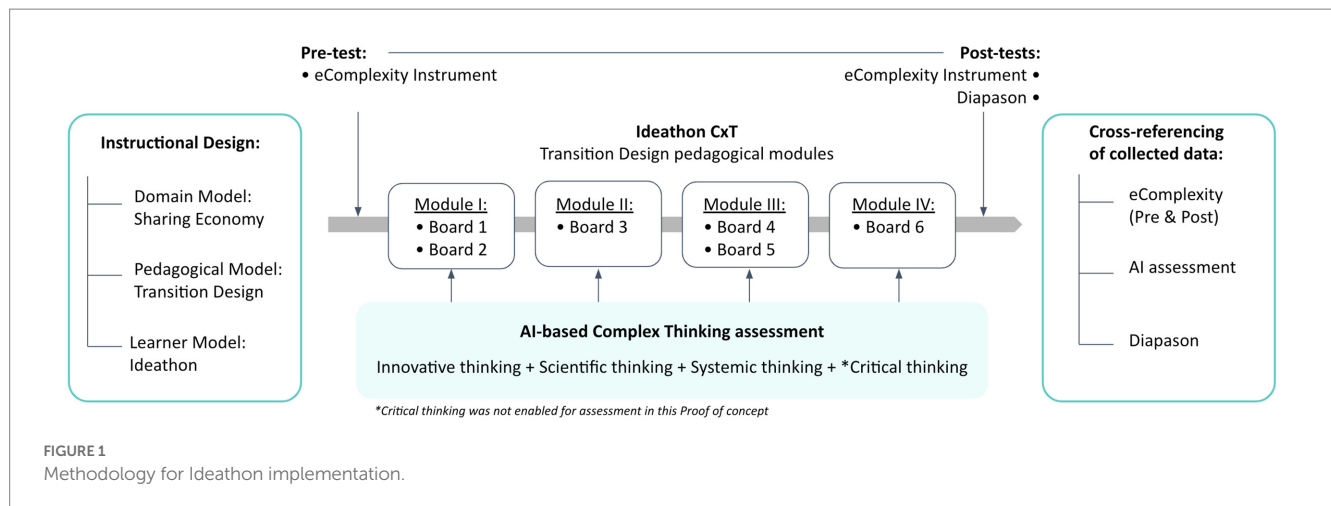
proposing synergistic solutions for the ideal future. Therefore, this methodology was chosen for the Ideathon event dynamics because it proceeds in a composite and logical way to analyze and search for solutions to a problem.

For the practical integration of Transition Design into the educational experience, we devised an idea-generation event based on the Ideathon concept. The Ideathon is a hackathon-style event in which a timed challenge is presented to be solved by the participants using innovative learning practices, usually in a competitive or collaborative environment (University of Washington, 2016; Barrow, 2021). The implementation of an Ideathon, which can be either face-to-face or virtual, has been applied to a variety of topics, including developing methodologies to enhance creativity (Yudina et al., 2021), preventing disasters through location-based gaming (Uesugi and Moriyama, 2020), supporting community development and revitalizing urban areas, and providing maker-driven solutions to regional strategic sector issues (SIC, 2021). The Ideathon implementation was based on module II of the program developed by Sanabria-Z et al. (2020), which integrates the following general recommendations: workshop rooms limited to 25 participants at tables with a maximum of 5 individuals; one instructor and one facilitator for mediation per room; 6 h of work including an opening and a closing conference, and a central screen displaying general instructions. Thus, while the Transition Design guides the participants in the operation of the on-screen activities, the Ideathon umbrella marks the big stages of the event to orchestrate the activities in the different rooms or venues.

4.1. Methodology for the Ideathon implementation

A total of 31 students participated in the Ideathon. As for the participant profile, it comprised undergraduate students between 18 and 23 from different disciplinary areas. They belonged to five universities in Mexico: Tecnológico de Monterrey (ITESM), which participated with the Mexico city (ITESM-CCM) and Guadalajara (ITESM-GDL) campuses; Universidad Autónoma de Ciudad Juárez (UACJ), in Chihuahua city; Instituto de Investigación, Innovación y Estudios de Posgrado para la Educación del Estado de Nuevo León (IIIEPE), in Monterrey city, and the Instituto Politécnico Nacional (IPN), in Mexico city. To manage the event, instructors received and briefed the participants, and facilitators assisted in logistical and technical matters only; neither were allowed to provide support in content issues. Regarding ethical considerations, students were asked to authorize the use of their data for research purposes by clicking on a button when registering on the platform. The Ideathon was run simultaneously in the five venues which were streaming online video from the classrooms throughout the event. Figure 1 shows the methodology followed for the implementation of the Ideathon.

Figure 1 shows the establishment of the Ideathon, which setting was based on Sanabria-Z et al. (2020). The starting point was the provision of an instructional design guideline that contemplated three types of models according to Hamal et al. (2022): domain model (sharing economy), pedagogical model (Transition Design) and learner model (Ideathon). It can be seen that four modules were integrated into the platform for implementation. Prior to the event, students were asked to answer the eComplexity instrument (Castillo-Martínez et al., 2022) to assess their perception of their level of



mastery of complex thinking, and also were provided with the sharing economy case study to get familiar with the topic. The day of the Ideathon, each student was assigned a specific theme to be addressed as a problem during the event, namely fashion, education, food industry, or health. Then, they proceeded to work through the 4 modules based on the Transition Design (Irwin, 2015) on the interactive platform as described below.

- a) Module 1, “Long-term vision,” consisted of boards 1 and 2. For the work on board 1 the participants started by placing an issue in a box related to the thematic problem in each of the 5 categories of social, economic, environmental, political and infrastructure. They reflected on possible connections or ramifications among all categories, aiming at making a full loop around all the categories until they could understand how the problem branched into all the areas. By using connecting lines between boxes they related the available issues to the original topic. On board 2 students mapped positive and negative relationships between stakeholders around the topic.

The main complex thinking subcompetency addressed in this module was systemic thinking.

- b) Module 2, “transition pathway,” consisted of board 3. Here, participants placed key issues drawn from board 1, then expanded them through three socio-technical levels according to the Transition Design approach: (1) large systems influences, (2) the “stuck” status quo, and (3) micro systems influences. They identified several themes or events at each of the three levels and their interconnections. Whenever they placed text in one of the levels, they were prompted to reflect on the potential connection, cause or what it led to in the other two levels.

In this module, systemic and critical thinking were the complex thinking subcompetencies addressed.

- c) Module 3, “medium-term visions” consisted of boards 4 and 5. For board 4 the students developed 3 milestones along the “transition path” from the present to the desired future. They tried to imagine what the situation would be like just before the desired long-term future and described it in narrative form. Then they did the same with the near future: what would a first step toward the long-term future

look like? This exercise taught them to think rigorously about transitions over long periods of time. They formed a narrative about a long-term transition from a problematic present to a desirable/sustainable long-term future (year 2100). For board 5, participants described 3 to 4 aspects of the vision for the future in each of the domains of daily life. The domains are a way of thinking at the levels of “organic” and nested systems from which everyday life emerges. The categories used to map the problem were set for mapping it, while the domains were aimed at encouraging them to think in a more integrated way about how to reconceive entire lifestyles to be more sustainable and place-based. Students thought of several facets of what daily life would be like at each level if the problem had been solved.

Three subcompetencies were addressed in this module: systemic, critical and scientific thinking.

- d) Module 4, “present,” consisted of board 6. Participants looked for “fragments” or “aspects” of the long-term vision of the future that may already be here, in the present, and listed them in cloud bubbles on the far right of the board. These “fragments of the future” are used as possible basis for systems interventions. They then developed concepts for potential “interventions” (solutions). They tried to place the interventions, in different areas of a matrix, in the area they wanted to change. The matrix helped them to place different types of interventions at different levels of scale within the social, political, economical, environmental, and technological dimensions, in cross-reference to the contextual nested systems levels consisting of household, neighborhood, city, state/region, nation and planet. Each intervention had to be connected to each other and to a milestone or long-term vision.

Figure 1 also shows the subcompetencies that were considered to be assessed by the AI platform, systemic thinking, scientific thinking and innovative thinking, however, critical thinking was not taken into account for this particular proof of concept test. After completion of the 4 modules, students were asked to answer two types of survey questionnaires, the eComplexity post-test, linked to the pre-test applied, and the Diapason (Alemán de la Garza, 2019) post-test, related to their perception of the interactive experience with the platform. The set of results from the eComplexity pre-and post-tests, the Diapason, and the complex thinking assessment provided by the AI platform were analyzed by examining their different crossovers and mutual influences.

TABLE 1 Three levels of complex thinking criteria for the IA platform.

Basic	Intermediate	Advanced
Innovative thinking		
The participant contributed 1 to 17 ideas in modules 1, 2, 3 (board 5) and 4.	The participant contributed 18 to 36 ideas in modules 1, 2, 3 (board 5) and 4.	The participant contributed more than 37 ideas in modules 1, 2, 3 (board 5) and 4.
Scientific thinking		
Match 1 to 3 terms along the boards of all modules.	Match 4 to 6 terms along the boards of all modules.	More than 7 terms coincide across the boards of all modules.
Systemic thinking		
The participant is able to establish 1 to 4 connections in the maps of modules 1 and 2.	The participant is able to establish 5 to 9 connections in the maps of modules 1 and 2.	The participant is able to establish more than 10 connections in the maps of modules 1 and 2.
Critical thinking		
The participant is able to make more elaborate proposals using a Boolean operator of the type (and, or, or not).	The participant is able to make more elaborate proposals using two to three Boolean operators of the type (and, or, or not).	The participant is able to make more elaborate proposals using more than three Boolean operators of the type (and, or, or not).

To evaluate participants' complex thinking meta-competency through AI, we created basic, intermediate, and advanced-level criteria for each sub-competency: systemic, scientific, critical, and innovative thinking. The criteria implied, for instance, that a participant digitally connects a content (box with words) with another content; or that they develop sentences with a considered construction elaborated by using operators between words (see [Table 1](#)).

4.2. Adapted works for the building of the Ideathon

A search for articles related to the subject of the present study was carried out. The search string used was as follows: TITLE-ABS-KEY (platform AND "artificial intelligence" AND competenc*). The inclusion criteria established were that the articles should be open access, that they should have been published in the period from 2019 to 2023 and that the type of document should correspond to articles, conferences, book chapters or books. Forty-seven documents were identified. Documents related to the health area ([Lamberti et al., 2019](#); [McNamara et al., 2019](#); [Shorey et al., 2019](#); [Rajadhyaksha, 2020](#); [Zhao et al., 2020](#); [Cheng et al., 2022](#); [Creed et al., 2022](#); [Liaw et al., 2022](#); [Lokala et al., 2022](#); [Shah et al., 2023](#)), industry ([Barykin et al., 2020](#); [Mokhtarname et al., 2020](#); [Sandner et al., 2020](#); [Dmitrievsky et al., 2022](#); [Obermayer et al., 2022](#); [Zakharkina et al., 2022](#)), education ([Hrich et al., 2019](#); [Tsalapatas et al., 2019](#); [Cortés et al., 2020](#); [Paba-Medina et al., 2020](#); [Raj et al., 2020](#); [Yang et al., 2020](#); [Demchenko et al., 2021](#); [Hurajová, 2021](#); [Jiang, 2021](#); [Petrescu et al., 2021](#); [Ghnemat et al., 2022](#); [Polak et al., 2022](#); [Ramírez-Montoya et al., 2022](#); [Rataj and Wojcik, 2022](#)), science ([Bruneckiene et al., 2019](#); [Desnos et al., 2022](#); [Ramírez-Montoya et al., 2022](#); [Zhu et al., 2022](#)), evaluation

([Kiran et al., 2019](#); [Prom et al., 2019](#); [Konys, 2020](#); [Bachiri and Mouncif, 2023](#); [Rashidi Fathabadi et al., 2023](#)), engineering ([Kaspar and Vielhaber, 2019](#); [Telnov and Korovin, 2019](#)), geography ([Abd Alsammed, 2022](#)), social problems ([Tubaro, 2022](#)), e-recruitment support ([Aljuaid and Abbod, 2020](#); [Krasovskiy et al., 2020](#)), ethics ([Hauer, 2022](#)) were found. It was possible to establish that there is no study on the use of platforms with AI to measure the competency of complex thinking, which is why this research is valuable.

The establishment of the Ideathon was framed by several studies. The studies that served as the basis for determining the use of the Transition Design technique were those of [Irwin, \(2015\)](#), [Juri et al. \(2021\)](#), [Bosch-Gómez et al. \(2022\)](#), [Cowart and Maione \(2022\)](#), [Irwin et al. \(2022\)](#) and [Sanabria-Zepeda and Santana-Castellón \(2022\)](#). The studies that served as guidance for the choice of an Ideathon to carry out the proof of concept of the platform with AI were those of the [University of Washington \(2016\)](#), [Sanabria-Z et al. \(2020\)](#), [Uesugi and Moriyama \(2020\)](#), [Barrow \(2021\)](#), [SIC \(2021\)](#) and [Yudina et al. \(2021\)](#) and the latter was key for the implementation because the recommendations regarding space, furniture, duration and the number of facilitators and instructors per mediation and room were considered. On the other hand, the studies by [Castillo-Martínez et al. \(2022\)](#) and [Alemán de la Garza \(2019\)](#) were considered because they address the design and validation process of the instruments that were applied for the present study, the first corresponds to the eComplexity instrument to measure the perception of students regarding their level of mastery of the complex thinking competency and the second corresponds to the Diapason instrument, which allows measuring the perception of the interactive experience with the platform.

5. Results

The general results of the Ideathon proof of concept are presented below in three sections: (5.1) experience of the Ideathon educational situation, (5.2) perception of complex thinking via eComplexity instrument, and (5.3) AI-based platform performance to measure complex thinking traits. The theoretical justification for the analysis of results, general outcomes of the proof of concept, and specific results observed in the event are presented below.

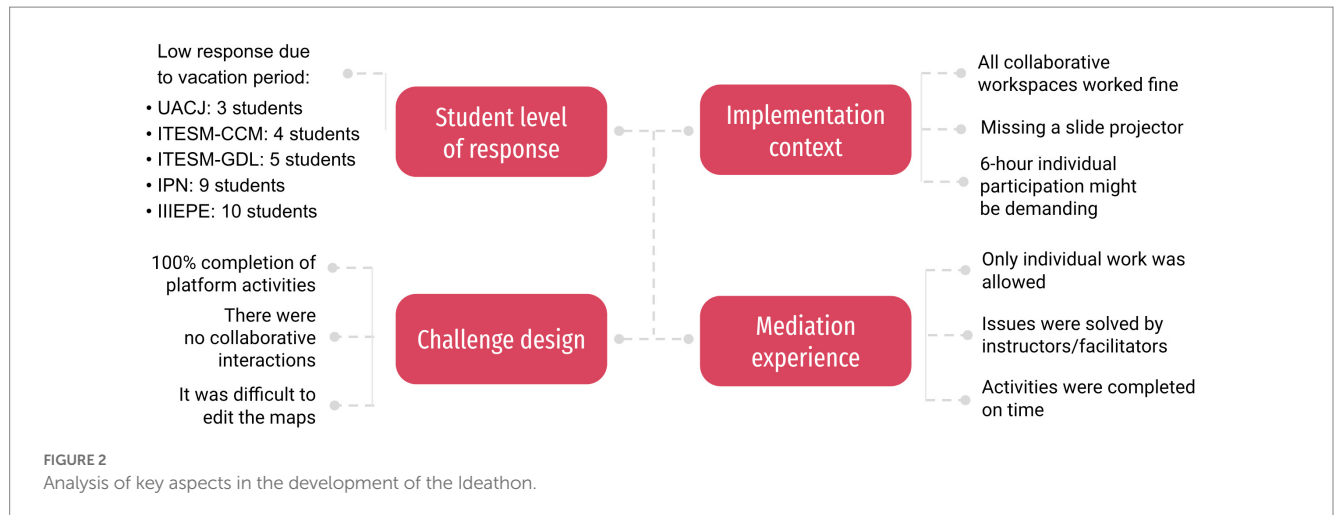
5.1. Experience of the Ideathon educational situation

The implementation of the Ideathon using the Transition Design technique with the theme of sharing economy allowed us to assess the following four aspects regarding students experience: (1) level of response, (2) implementation context, (3) mediation experience, and (4) challenge design.

[Figure 2](#) shows the four aspects considered above in relation to students participation and experience in Ideathon.

5.1.1. Student level of response

When it comes to planning an event, timing is often a critical point that can greatly affect an occasion if it intersects with a school schedule of calendar events such as vacations, exams and academic events ([University of Waterloo, n.d.](#)). According to [Ljubisic \(2017\)](#), some motivations for students to attend college related events include



the possibility of socialization in a light atmosphere, be provided with food and beverage, listen to guest speakers, and participating in thought-provoking workshops. Regarding these aspects, the conditions at Ideathon were as follows. Students were either directly invited by their teachers or recommended by other teachers to participate. Each setting was different, some were traditional formal classrooms while others were informal settings (e.g., design thinking classroom), but the atmosphere was relaxed in all settings since this was a non curricular activity, and some campus were already in winter vacations. Food and non-alcoholic drinks were provided in each location (e.g., pizza and soda). Although there was no speaker *per se*, the principal investigator gave a welcome address and explained the significance of the Ideathon. The event was promoted through social networks with a poster encouraging participation with the title “Ideatón: future of sharing economy,” including the key themes of Transition Design, AI, sharing economy and complex reasoning.

As for the specific results related to this aspect, timing was certainly a major drawback that yielded little student participation because two or the four participant universities were already on winter vacation. A total of 31 students participated in the event, both the IIEPE (10 students) and the IPN (9) had more participation because the instructors were still having sessions with their students. Moreover, the day of the event also turned out to be a special occasion for administrative acts in three venues, UACJ (3), ITESM-CCM (4) and ITESM-GDL (5). Considering that each location had the capacity to receive about 20 students in each classroom for a potential grand total of 100, the attendance of 31 students was considered particularly low. Among the recommendations for future studies is to ensure that the times in which activities are carried out as part of the AI platform are optimal for greater participation.

5.1.2. Implementation context

De Pretto et al. (2019) mention the importance of classroom design, facilities and conditions to stimulate the learning experience in higher education, where improved equipment, arrangement flexibility, attractive decoration, and adding natural elements play a key role. According to these aspects, the configuration was as follows. Regarding equipment, all sites provided Internet connection via Wi-Fi or Ethernet cable. Some campuses provided desktop computers while others asked students to bring their laptops or tablets. As for the

classroom configuration, the distribution of the desks varied by venue, including semicircular, square, linear and double line arrangements. Almost all classrooms provided a main screen where the main and other classrooms’ broadcasts were projected and heard. The interior design was also varied by venue, including formal classrooms with desks covered with tablecloths, robotics classrooms surrounded with electronic accessories, and design thinking classrooms with inflatables hanging from the ceiling. No natural elements were identified in the classrooms, except that some were surrounded by large windows while others were closed classrooms.

As for the particular results in this area, the UACJ could not provide a space for the event, so they had to resort to using a classroom at the Tecnológico de Monterrey, Ciudad Juárez campus, while in the ITESM-CCM students use the Robotics Laboratory because of graduation events, as noted by their instructor: “We lacked a monitor to project the synchronous participation from the other venues, which made it difficult to listen to the instructions and hold the participants’ attention.” Despite the logistical complications, the overall spaces, furniture, and equipment were adequate for the students to complete on time all the activities on the agenda.

5.1.3. Challenge design

According to Ifenthaler et al. (2018), research on the employment of challenges in online learning environments emphasizes their relationship to learning performance, where the number of activities started and completed in a challenge-based online platform are the most reliable predictor of student learning performance, also linked to the individual investment of time and effort. In terms of the AI-based platform used in the Ideathon, its instructional design included six activities that had to be completed in sequence in order to complete the challenge. Each activity was represented on a board that included instructions and examples of how to respond to the specific challenge using text and digital elements (e.g., connecting lines and boxes). In order to advance from one board to the next, it was necessary for the student to save the achievement of the section, and press the button to continue to the next section.

In terms of the specific results of this aspect, the following is observed. Of the total number of registered participants, not all completed the pre-test or post-test questionnaires, and not all

showed up for the event. However, as far as fulfilling the platform activities is concerned, all students went through the 6-board sequence, which translates into 100% of the participants successfully completing the sharing economy challenge. The takeaway is that both the instructional design and performance of the platform functioned according to the expectations of the proof of concept. Regarding the time an effort, the configuration of the platform was designed for individual work, which was adequate according to expectations, however, it is acknowledged that 6 hours of ideathon on an individual basis can be demanding. Also, the possibility of collaborative work and its effects on the development of competencies were not factored in this proof of concept. Furthermore, some difficulties in using the AI platform were reported as part of the interactive experience. One of the instructors stated this issue as: “The platform did not save its maps. It was difficult to edit the maps because they were out of configuration. It was not possible to add colors to the rectangles. It was not possible to use different types of figures to make the maps.” Although this was not necessarily a common occurrence, there were several different difficulties in using the AI platform that were collected through the Diapason instrument that should make it possible to establish further improvements regarding usability and functionality.

5.1.4. Mediation experience

Distinctive features of mediation in hybrid single events such as the Ideathon are similar to those that apply to online workshops. According to the German management software company SAP (Steinmetz, 2022, November 9), best practices for digital workshop mediation include frequent contact with participants, overcommunicating, and designing for maximum engagement. The design of the Ideathon experience integrated the participation per classroom of an instructor, in charge of providing instructions and time supervision, and a facilitator, in charge of taking care of logistical contingencies, while both could mediate regarding the use of the platform and technical problems. However, none should provide assistance on questions related to the content of the challenges. Each desk was intended to have a maximum of 4 students, each working individually to solve a challenge on a different topic from that of the neighboring students. They were allowed to talk about topics not related to the activity, but in case they had any related questions, they could approach the instructor or facilitator.

In this context, the specific results were as follows. General directions were given by the main researcher during an introductory explanation broadcasted to all the participants who were in the different venues. Instructors and facilitators properly played their role as observers of the conduct and performance of the participants, guiding the students if questions arose, but taking care that they could advance independently by following the instructions that were integrated into the platform. Although the physical facilities may well allow for collaborative work, the AI platform did not yet provide the algorithm possibilities to measure team performance; therefore, instructors and facilitators supervised that only individual work was conducted during the event. Furthermore, the event design and mediation appeared to contribute to engaging the students since they were able to overcome all stages according to the time limits.

5.2. Perception of complex thinking via eComplexity instrument

Students' perception of their competency in complex thinking was conducted by applying the eComplexity instrument before and after the Ideathon. The eComplexity instrument (Ramírez-Montoya et al., 2022) is a five-point Likert-type scale questionnaire: do not agree at all (1), slightly agree (2), neither agree nor disagree (3), agree (4) and strongly agree (5). The instrument consisted of the following indicators: Knowledge, skills and attitudes or values for the four dimensions of innovative thinking, scientific thinking, systemic thinking, and critical thinking, which were integrated into 25 items. It was applied as a pre-test and post-test to identify whether the participants' perceptions had changed when they finished the four blocks of the Ideathon.

Table 2 shows the mean scores obtained before and after the Ideathon.

As shown in Table 2, although 31 students participated in the Ideathon, only 18 students answered both the pre-test and post-test eComplexity instrument. The Student's *t*-test performed on the 18 cases indicates that the means obtained in the pre-test and post-test had statistically significant differences, implying that the results are

TABLE 2 Pre test and post test means of eComplexity questionnaire.

Type of thinking	Item	Mean pre test	Mean post test
Critical thinking	14	4.47	4.35
	15	4.00	4.24
	16	4.29	4.41
	17	4.18	4.41
	18	4.06	4.29
	19	4.71	4.53
Innovative thinking	20	3.94	4.18
	21	4.00	4.41
	22	3.82	4.12
	23	3.94	4.29
	24	4.41	4.35
	25	4.41	4.41
Scientific thinking	7	3.94	4.06
	8	4.00	4.06
	9	4.18	4.18
	10	4.00	4.12
	11	4.00	4.18
	12	4.00	4.29
	13	4.35	4.35
Systemic thinking	1	4.18	4.29
	2	4.47	4.41
	3	4.24	4.41
	4	4.41	4.24
	5	4.82	4.65
	6	4.41	4.41

reliably different. The result of the Student's *t*-test was $t(18)=2.06$; $p=0.008$. When performing an analysis by types of thinking, we observed that the mean for scientific thinking increased in the post-test, indicating that the students perceived a higher level of mastery in this sub-competency.

5.3. AI-based platform performance to measure complex thinking traits

The AI-based platform was enabled to identify features of the complex thinking competency, which could complement the results of the students' own perception of complex thinking from the eComplexity instrument. To this end, a rubric was used with three levels of mastery of complex thinking, basic, intermediate and advanced (see Table 1), which were programmed into the platform through identifying student interaction behaviors with the platform as well as the characteristics of the texts they inserted. The rubric was then transformed into an algorithm which, by analyzing the boards using a decision tree classifier (Jijo and Abdulazeez, 2021), was able to extract from each board the number of ideas, logical conjunctions, and existing connections; this was done by converting the boards into JSON format files, so that all the contents can be read as text strings. Of the four sub-competencies of complex thinking, this proof of concept test focused on three, innovative, scientific, and systemic thinking, leaving the integration of critical thinking to be tested in a future edition. Specifically, complex thinking traits were calculated according to the following programming on the platform. The innovative thinking trait was measured by calculating the number of ideas inserted throughout the 4 blocks, where the intermediate level is considered to be the range between 18 and 36 ideas along all boards; an idea was equivalent to a sentence placed in a box, which was identified by the system by tracking what was written between periods. For the scientific thinking trait, terms were identified that appeared in more than one board across the four blocks, where the intermediate level is considered to be the range between 4 to 6 terms along all boards; a term consisted of a word or construct, which was identified by the system as distinct from pronouns and articles. The systemic thinking trait was addressed through identifying the number of connections made in the maps, along blocks 1 and 2, where the intermediate level is considered to be the range 5 to 9 connections in both blocks; the connections were equivalent to the lines made between boxes by the students, which the system identified as a single line (and not the use of arrows or other connecting elements).

The results of the mastery levels are shown in Table 3.

Table 3 shows that concerning scientific and systemic thinking, there was a clear predominance at the advanced level with 23 students

on each, contrary to innovative thinking where only 3 students reached the advanced level. However, the majority of participants achieved an intermediate level in the innovative thinking trait, with a total of 26 students.

6. Discussion

Carrying out learning activities through an Ideathon using the Transition Design approach under a time limit encouraged participants to generate innovative solutions to complex problems. Figure 1 showed the challenges that arose before and during the Ideathon: low response from students due to the winter vacation and difficulty in designing the maps requested in the Transition Design activities with the theme of sharing economy; however, despite the difficulties, all participants managed to complete the agenda activities on time. Against the unexpected in the process, following the recommendations by Sanabria-Z et al. (2020) regarding what should be taken care of when carrying out an Ideathon contributed to the successful completion of the activities (e.g., duration, numbers of students per table). Furthermore, the overall Ideathon setting was in line with what Haro (2018) mentioned regarding the role of the Ideathon: promoting a culture of innovation was achieved by using 4.0 technology, the event was held simultaneously in different venues, and the products generated by the participants were shared online. Combining the Ideathon with Transition Design to achieve a learning target was an optimal pedagogical formula that can be replicated using other megatrends as a central theme, as was the case with the sharing economy.

Applying the eComplexity instrument makes it possible to identify changes in the student's perception of mastery of the complex thinking meta-competency after an Ideathon-style intervention. Table 2 shows that changes in students' perception of mastery occurred, for instance, the high increase in the scientific thinking sub-competency. The effectiveness of measuring complex thinking sub-competencies in the context of the sharing economy is consistent with the results of the study by Vázquez-Parra et al. (2022). The application of the eComplexity instrument allows us to know precisely whether an intervention with university students improves students' perception of the level of mastery of their complex thinking competency.

The tested IA platform goes beyond measurement in perceptual terms, identifying levels of participants' mastery of complex thinking sub-competencies. Table 3 shows that most participants were at an intermediate level of mastery of innovative thinking, contrary to the systemic and scientific types of thinking, in which they were at an advanced level. AI has been used to monitor and track student interactions, showing enormous potential to improve learning, teaching, assessment, and educational administration (Xia et al., 2022). Through the proof of concept, it was possible to identify that an AI platform has the potential to measure the levels of mastery of the complex thinking meta-competency and can serve as a basis for the design of strategies that strengthen its sub-competencies.

Using a variety of instruments to measure complex thinking competency allows for a more robust assessment outcome. The results regarding the perception of the participants thanks to the application of the eComplexity instrument are important (Table 2), but it is valuable to be able to have intersected results that allow us to know practical

TABLE 3 Complex thinking subcompetencies' levels of mastery through the AI platform.

Level of mastery	Innovative thinking	Scientific thinking	Systemic thinking
Basic	2	4	2
Intermediate	26	4	6
Advanced	3	23	23
Total	31	31	31

levels of mastery through the measurement through the platform with the use of AI (Table 3). Xia et al. (2022) point out that assessment is one of the educational domains where there is potential for the use of AI. The analysis of instrument outputs in the process of interpreting the results showed that there was an increase in student perceptual terms with respect to their level of mastery of the four sub-competencies of complex thinking (i.e., critical, innovative, scientific, and systemic thinking), which was in accordance with the results for the three complex thinking sub-competency traits identified by the AI-based platform (i.e., innovative, scientific, and systemic thinking), since most participants achieved intermediate (innovative thinking) and advanced (systemic and scientific thinking) levels of mastery.

7. Conclusion

The stated hypotheses were validated for the development of complex thinking. H1, regarding the event as a pedagogical situation, was true because the Ideathon and the guide of the Transition Design achieved the expected performance in the allotted time. H2, related to the use of AI for measuring complex thinking traits, also proved true because the criteria created for each sub-competency allowed us to assign basic, intermediate, or advanced values for the mastery level of complex thinking.

The implications for practice are that an implementation model has been produced that can be replicated to test different themes beyond the sharing economy megatrend. This combination of physical (Ideathon) and digital (Transition Design) environment gives us a guideline to generate research on hybrid achievement in pedagogy in an accelerated and somewhat competition-driven fashion. Likewise, the exercise of making a proof of concept using AI to measure mastery of complex thinking is a pedagogical experience that can be transposed to different subjects of study and even be incorporated into LMS platforms that are used daily.

The present research focused on a first proof of concept. We note limitations in the quality of the graphics and usability of the platform; the lack of collaborative activities that can be integrated with individual ones to be measured by the platform; the testing of different event lengths to measure performance over days, weeks, or months; the sparse content entered into the platform by students which does not yet allow AI to flourish, for example, by replacing facilitators with a chat system for student queries. Future studies could test different event lengths to measure performance over days, weeks, or months. Other opportunities include using additional criteria to measure the developmental behaviors of complex thinking, its traits, and its sub-competencies in greater detail; monitoring other types of competencies such as problem-solving, computational thinking, or collaboration; and testing functional interactive aspects that are

relevant to make the use of the platform more user-friendly, functional and attractive.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Regarding ethical considerations, students were asked to authorize the use of their data for research purposes by clicking on a button when registering on the platform.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

The authors acknowledge the financial support of Tecnológico de Monterrey through the “Challenge-Based Research Funding Program 2022” (Project ID # I004-IFE001-C2-T3-T). The authors also wish to acknowledge the financial and technical support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, in the production of this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abd Alsammed, S. M. Z. (2022). Advanced GIS-based multi-function support system for identifying the best route. *Baghdad Sci. J.* 19:631. doi: 10.21123/bsj.2022.19.3.0631
- Acquier, A., Daudigeos, T., and Pinkse, J. (2017). Promises and paradoxes of the sharing economy: an organizing framework. *Technol. Forecast. Soc. Chang.* 125, 1–10. doi: 10.1016/j.techfore.2017.07.006
- Akmeşe, Ö. F., Kör, H., and Erbay, H. (2021). Use of machine learning techniques for the forecast of student achievement in higher education. *Inf. Technol. Learn. Tools* 82, 297–311. doi: 10.33407/itlt.v82i2.4178
- Alemán de la Garza, L. (2019). Modelo de indicadores de calidad para cursos en-línea, masivos y abiertos (mooc). Caso de un MOOC para el desarrollo profesional docente. Ciudad de México, México: Editora Nómada.

- Alhumaid, K., Naqbi, S., Elsoori, D., and Mansoori, M. (2023). The adoption of artificial intelligence applications in education. *Int. J. Data Netw. Sci.* 7, 457–466. doi: 10.5267/j.ijdns.2022.8.013
- Aljuaid, A., and Abbod, M. "Artificial intelligence-based E-recruitments system," (2020) 2020 IEEE 10th International Conference on Intelligent Systems (IS), Varna, Bulgaria, 144–147.
- Alsaleh, N. J. (2020). Teaching critical thinking skills literature review. *Turkish Online J. Educ. Technol.* 19, 21–39.
- Bachiri, Y.-A., and Mouncif, H. (2023). Artificial intelligence system in aid of pedagogical engineering for knowledge assessment on MOOC platforms: open EdX and moodle. *Int. J. Emerg. Technol. Learn.* 18, 144–160. doi: 10.3991/ijet.v18i05.36589
- Barile, S., Orecchini, F., Saviano, M., and Farioli, F. (2018). People, technology, and governance for sustainability: the contribution of systems and cyber-systemic thinking. *Sustain. Sci.* 13, 1197–1208. doi: 10.1007/s11625-018-0621-y
- Bäro, A., Toepler, F., Meynhardt, T., and Velamuri, V. K. (2022). Participating in the sharing economy: the role of individual characteristics. *Manag. Decis. Econ.* 43, 3715–3735. doi: 10.1002/mde.3624
- Barrow, J. (2021). Hack 2040: an Ideathon event for engaging students in university policy while enhancing their enterprise skills. In Hack 2040
- Barykin, S., Borovkov, A., Rozhdvestvenskiy, O., Tarshin, A., and Yadykin, V. (2020). Staff competence and training for digital industry. *IOP Conf. Ser.: Mater. Sci. Eng.* 940:012106. doi: 10.1088/1757-899X/940/1/012106
- Bengu, A., Abrignani, E., Sabuncuoglu, I., and Yilmaz, C. (2020). Rethinking higher education for the emerging needs of society. *Glob. Solutions J.* 5, 178–187.
- Bonfield, C. A., Salter, M., Longmuir, A., Benson, M., and Adachi, C. (2020). Transformation or evolution?: Education 4.0, teaching and learning in the digital age. *Higher Educ. Pedagogies* 5, 223–246. doi: 10.1080/23752696.2020.1816847
- Bosch-Gómez, S., Ortega-Pallanez, M., and Dorn, E. (2022). Diseño para la Transición: Entrelazando saberes por medio de la conversación. Cuadernos del Centro de Estudios en Diseño y Comunicación. *Ensayos* 157, 65–86. doi: 10.18682/cdc.vi157.6845
- Bruneckiene, J., Robertas, J., Ineta, Z., Jonas, R., and Mantas, L. (2019). Assessment of investment attractiveness in European countries by artificial neural networks: what competences are needed to make a decision on collective well-being? *Sustainability* 11:6892. doi: 10.3390/su11246892
- Castillo-Martínez, I. M., and Ramírez-Montoya, M. S. (2022). Instrumento eComplexity: Medición de la percepción de estudiantes de educación superior acerca de su competencia de razonamiento para la complejidad. <https://hdl.handle.net/11285/643622>
- Chen, G. (2021). Design of cooperation platform based on HTML5 framework. In International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy. 233–238. Springer, Cham.
- Cheng, W., Zhang, Z., Hoelzer, S., Tang, W., Liang, Y., Du, Y., et al. (2022). Evaluation of a village-based digital health kiosks program: A protocol for a cluster randomized clinical trial. *Digital Health*. 8. doi: 10.1177/20552076221129100
- Chiu, T. K. (2021). A holistic approach to the design of artificial intelligence (AI) education for K-12 schools. *TechTrends* 65, 796–807. doi: 10.1007/s11528-021-00637-1
- Cornejo-Velazquez, E., Clavel-Maqueda, M., Perez-Lopez-Portillo, H., and Lyubimova, E. (2020). Business model of learning platforms in sharing economy. *Electron. J. e-Learn.* 18, 102–101. doi: 10.34190/EJEL.20.18.1.008
- Cortés, D., Molina, A., and Ramírez, J. (2020). Open innovation laboratory: Education 4.0 environments to improve competences in scholars. Proceedings of the 18th LACCEI International Multi-Conference for Engineering, Education and Technology.
- Costa-Mendes, R., Oliveira, T., Castelli, M., and Cruz-Jesus, F. (2021). A machine learning approximation of the 2015 Portuguese high school student grades: a hybrid approach. *Educ. Inf. Technol.* 26, 1527–1547. doi: 10.1007/s10639-020-10316-y
- Cowart, A., and Maione, D. (2022). Transitioning toward the slow and long: developing experiential futures approach toward system change in fashion. Cuadernos del Centro de Estudios en Diseño y Comunicación. *Ensayos* 157, 29–63. doi: 10.18682/cdc.vi157.6844
- Creed, T. A., Salama, L., Slevin, R., Tanana, M., Imel, Z., Narayanan, S., et al. (2022). Enhancing the quality of cognitive behavioral therapy in community mental health through artificial intelligence generated fidelity feedback (Project AFFECT): a study protocol. *BMC Health Serv. Res.* 22:1177. doi: 10.1186/s12913-022-08519-9
- Demchenko, M. V., Gulieva, M. E., Larina, T. V., and Simaeva, E. P. (2021). Digital transformation of legal education: problems, risks and prospects. *Eur. J. Contemp. Educ.* 10, 297–307. doi: 10.13187/ejced.2021.2.297
- Desnos, K., Bourgoignie, T., Dardaillon, M., Sourbier, N., Gesny, O., and Pelcat, M. (2022). Ultra-fast machine learning inference through C code generation for tangled program graphs. *IEEE Workshop on Signal Processing Systems (SiPS)*, 1–6.
- De Pretto, L., Hii, C. L., Chiang, C. L., Ong, S. P., Tan, D. E. S. Y., and Pang, C. H. (2019). On the key factors in higher education classroom design: Physical aspects of the undergraduate teaching and learning experience in Malaysia. doi: 10.20944/PREPRINTS201908.0089.V1
- Di Bella, D. (2018). "Impacto de la experiencia Diseño en Perspectiva" in *Cuadernos del Centro de Estudios en Diseño y Comunicación [Ensayos]*, vol. 80 (Buenos Aires: Facultad Diseño y Comunicación. Universidad de Palermo), 173–219.
- Di Bella, D. V. (2022). Experiencia Diseño en Perspectiva: Diseño para la Transición 2° Informe de Impacto (2018–2022). *Cuadernos del Centro de Estudios de Diseño y Comunicación* 158. doi: 10.18682/cdc.vi158.6954
- Dmitrievsky, A. N., Eremin, N. A., Safarova, E. A., and Stolyarov, V. E. (2022). Implementation of complex scientific and technical programs at the late stages of operation of oil and gas fields. *SOCAR Proc.* 2022, 1–8. doi: 10.5510/OGP2022SI200728
- Ghnem, R., Shaout, A., and Al-Sow, A. M. (2022). Higher education transformation for artificial intelligence revolution: transformation framework. *Int. J. Emerg. Technol. Learn.* 17, 224–241. doi: 10.3991/ijet.v17i19.33309
- Giunta, B., Burnay, C., Maiden, N., and Faulkner, S. (2022). Creativity triggers: extension and empirical evaluation of their effectiveness during requirements elicitation. *J. Syst. Softw.* 191:111365. doi: 10.1016/j.jss.2022.111365
- Hamal, O., El Faddouli, N. E., Harouni, M. H. A., and Lu, J. (2022). Artificial intelligent in education. *Sustainability* 14:2862. doi: 10.3390/su14052862
- Haro, C. (2018). Festival de Innovación Epicentro. *Soc. Innov. J.*. Available at: <https://socialinnovationsjournal.org/editions/issue-2011/75-RED>. Revista de Educación a Distancia. 62, 20, 01, 2020, 10.6018/red.398381
- Hauer, T. (2022). Incompleteness of moral choice and evolution towards fully autonomous AI. *Humanit. Soc. Sci. Commun.* 9, 1–10. doi: 10.1057/s41599-022-01060-4
- Hou, M. (2018). Construction and evaluation of a blending teaching model of linear algebra and probability statistics in the internet+ background by using the gradient Newton combination algorithm. In International Conference on Frontier Computing. 1525–1532. Springer, Singapore
- Hrich, N., Lazaar, M., and Khaldi, M. (2019). MAPSS an intelligent architecture for the pedagogical support: experiment, results & discussion. *Int. J. Emerg. Technol. Learn.* 14, 61–79. doi: 10.3991/ijet.v14i21.11012
- Hu, J. (2021). Teaching evaluation system by use of machine learning and artificial intelligence methods. *Int. J. Emerg. Technol. Learn.* 16, 87–101. doi: 10.3991/ijet.v16i05.20299
- Hurajová, L. (2021). Can close cooperation between ESP/CLIL experts and disciplinary teachers in higher education lead to fostering English education environment. *J. Teach. Eng. Specific Acad. Purp.* 9, 129–136. doi: 10.22190/JTESAP2101129H
- Ifenthaler, D., Gibson, D., and Zheng, L. (2018). "Attributes of engagement in challenge-based digital learning environments" in *15th International Conference on Cognition and Exploratory Learning in Digital Age CELDA 2018*. <https://eric.ed.gov/?id=ED600598>
- Irwin, T. (2012). *Wicked problems and the relationship triad* En grow small, think beautiful: ideas for a sustainable world from Schumacher College, Floris Books.
- Irwin, T. (2015). Transition design: a proposal for a new area of design practice, study, and research. *Des. Cult.* 7, 229–246. doi: 10.1080/17547075.2015.1051829
- Irwin, T., Tonkinwise, C., and Kossoff, G. (2022). Transition Design: an educational framework for advancing the study and design of sustainable transitions. *Cuadernos del Centro de Estudios de Diseño y Comunicación* 105, 31–72. doi: 10.18682/cdc.vi105.4188
- Ivanova, A. O., Ignatieva, T. A., and Pilyavsky, V. P. (2021). Trends, opportunities and perspectives of digital education development in the global economy. *IOP Conf. Ser.: Earth Environ. Sci.* 666:032033. doi: 10.1088/1755-1315/666/3/032033
- Jarke, J., and Macgilchrist, F. (2021). Dashboard stories: how narratives told by predictive analytics reconfigure roles, risk and sociality in education. *Big Data Soc.* 8:20539517211025561. doi: 10.1177/2053951721102556
- Jiang, L. (2021). Virtual reality action interactive teaching artificial intelligence education system. *Complexity* 2021, 1–11. doi: 10.1155/2021/5553211
- Jijo, B. T., and Abdulazeez, A. M. (2021). Classification based on decision tree algorithm for machine learning. *Evaluation* 6. doi: 10.38094/jastt20165
- Juri, S., Zurbriggen, C., Bosch Gómez, S., and Ortega Pallanez, M. (2021). Transition Design in Latin America: enabling collective learning and change. *Front. Sociol.* 6:725053. doi: 10.3389/fsoc.2021.725053
- Kaspar, J., and Vielhaber, M. (2019). Cooperative and application-oriented learning in engineering design—systems design methodology educated on solutions for mousetrap-powered vehicles. Proceedings of the 21st International Conference on Engineering and Product Design Education: Towards A New Innovation Landscape, E and PDE 2019
- Kiran, A., Butt, W. H., Anwar, M. W., Azam, F., and Maqbool, B. (2019). A comprehensive investigation of modern test suite optimization trends, tools and techniques. *IEEE Access* 7, 89093–89117. doi: 10.1109/ACCESS.2019.2926384
- Konys, A. (2020). How to support digital sustainability assessment? An attempt to knowledge systematization. *Procedia Comput. Sci.* 176, 2297–2311. doi: 10.1016/j.procs.2020.09.288

- Krasovskiy, I. N., Pilyavski, V. P., Shendrikova, S. P., and Nazrieva, M. V. (2020). Mechanism of the innovation development in the university. In *International Scientific Conference "Far East Con"* (ISCSEC 2020). 2297–2311. Atlantis Press
- Kuhn, D. (2002). "What is scientific thinking, and how does it develop?" in *Blackwell handbook of childhood cognitive development*. ed. U. Goswami (Oxford: Blackwell Publishing), 371–393.
- Lamberti, M. J., Wilkinson, M., Donzanti, B. A., Wohlhieter, G. E., Parikh, S., Wilkins, R. G., et al. (2019). A study on the application and use of artificial intelligence to support drug development. *Clin. Ther.* 41, 1414–1426. doi: 10.1016/j.clinthera.2019.05.018
- Liang, Y., and Wang, Y. (2020). "Integration of information technology and three-dimensional teaching mode" in *Data processing techniques and applications for cyber-physical systems (DPTA 2019). Advances in intelligent systems and computing*. eds. C. Huang, Y. W. Chan and N. Yen (Singapore: Springer), 1088.
- Liaw, W., Kueper, J. K., Lin, S., Bazemore, A., and Kakadiaris, I. (2022). Competencies for the use of artificial intelligence in primary care. *Ann. Fam. Med.* 20, 559–563. doi: 10.1370/afm.2887
- Ljubisic, I. (2017). Student Engagement in College Related Events. [Master's Thesis, Rochester Institute of Technology]. *RIT Scholar Works Repository*. <https://scholarworks.rit.edu/theses/10030/>
- Lokala, U., Lamy, F., Daniulaityte, R., Gaur, M., Gyrard, A., Thirunarayan, K., et al. (2022). Drug abuse ontology to harness web-based data for substance use epidemiology research: ontology development study. *JMIR Public Health Surveill.* 8:e24938. doi: 10.2196/24938
- Marks, A., and Al-Ali, M. (2022). Digital transformation in higher education: a framework for maturity assessment. In *COVID-19 challenges to university information technology governance* 61–81. Springer, Cham.
- McGinnis, W. D., Siu, C., Andre, S., and Huang, H. (2018). Category encoders: a Scikit-learn-contrib package of transformers for encoding categorical data. *J. Open Source Softw.* 3:501. doi: 10.21105/joss.00501
- McNamara, D. M., Goldberg, S. L., Latts, L., Atieh Graham, D. M., Waintraub, S. E., Norden, A. D., et al. (2019). Differential impact of cognitive computing augmented by real world evidence on novice and expert oncologists. *Cancer Med.* 8, 6578–6584. doi: 10.1002/cam4.2548
- Mokhtarname, R., Safavi, A. A., Urbas, L., Salimi, F., Zerfat, M. M., and Harasi, N. (2020). Toward HAZOP 4.0 approach for managing the complexities of the hazard and operability of an industrial polymerization reactor. *IFAC-PapersOnLine* 53, 13593–13600. doi: 10.1016/j.ifacol.2020.12.852
- Munir, A., and Awan, O. A. (2022). Creative thinking: its importance and how to cultivate it. *Acad. Radiol.* 29, 1610–1612. doi: 10.1016/j.acra.2021.06.001
- Muñoz, P., and Cohen, B. (2017). Mapping out the sharing economy: a configurational approach to sharing business modeling. *Technol. Forecast. Soc. Chang.* 125, 21–37. doi: 10.1080/02635143.2013.878328
- Obermayer, N., Nagy, V., Tóth, V. E., and Kiss, D. (2022). Overcoming the challenges of digitalisation in Hungarian manufacturing companies. *Proceedings of the European Conference on Knowledge Management, ECKM*, 23, 845–851
- Paba-Medina, M. C., Acosta-Medina, J. K., and Torres-Barreto, M. L. (2020). Design of a gamified tool for the development of citizenship competencies. *International Conference on Higher Education Advances*, 743–750
- Passig, D., and Cohen, L. (2014). Measuring the style of innovative thinking among engineering students. *Res. Sci. Technol. Educ.* 32, 56–77. doi: 10.1080/02635143.2013.878328
- Petrescu, D., Enache, D., and Dutu, L. (2021). Collaborative decision-making in online education. *Procedia Comput. Sci.* 199, 1090–1094. doi: 10.1016/j.procs.2022.01.138
- Polak, S., Schiavo, G., and Zancanaro, M. (2022). Teachers' Perspective on Artificial Intelligence Education: an Initial Investigation. *Conference on Human Factors in Computing Systems - Proceedings*. doi: 10.1145/3491101.3519866
- Prom, N. A., Litvinova, E. A., Shokhnekh, A. V., Yovanovich, T. G., and Lomakin, N. I. (2019). Perseptron for assessing the students' task performance in learning a foreign language according to the competencies using bi data. *IOP Conf. Ser.: Mater. Sci. Eng.* 483:012048. doi: 10.1088/1757-899X/483/1/012048
- Raj, R. K., Romanowski, C. J., Impagliazzo, J., Aly, S. G., Becker, B. A., Chen, J., et al. (2020). High performance computing education: current challenges and future directions. *Annual Conference on Innovation and Technology in Computer Science Education, ITICSE*, 51–74
- Rajadhyaksha, V. D. (2020). Medical affairs post-COVID 19: are we ready to take the baton? *Perspect. Clin. Res.* 11, 124–127. doi: 10.4103/picr.PICR_164_20
- Ramírez-Montoya, M. S., Castillo-Martínez, I. M., Sanabria-Z, J., and Miranda, J. (2022). Complex thinking in the framework of education 4.0 and open innovation—a systematic literature review. *J. Open Innov.: Technol., Market, Complex.* 8. doi: 10.3390/joitmc8100004
- Rashidi Fathabadi, F., Grantner, J. L., Shebrain, S. A., and Abdel-Qader, I. (2023). 3D autonomous surgeon's hand movement assessment using a cascaded fuzzy supervisor in multi-thread video processing. *Sensors* 23. doi: 10.3390/s23052623
- Rataj, M., and Wojcik, J. (2022). The impact of COVID-19 on online final exam scores among computer science students. *Proceedings of the European Conference on Knowledge Management, ECKM*, 23, 955–964
- Sanabria-Z, J., Davidson, A.-L., Romero, M., and Quintana, T. (2020). Macro-dissemination of maker cultures: 21st century competencies through an Ideathon. *Rev. Educ. Distancia*. 20. doi: 10.6018/red.398381
- Sanabria-Zepeda, J., and Santana-Castellón, E. (2022). TD-HUMETAV: Análisis de problemas perversos para el Museo de Ciencias Ambientales, con base en el Diseño para la Transición. *Cuadernos del Centro de Estudios de Diseño y Comunicación* 157. doi: 10.18682/cdc.vi157.6851
- Sandner, P., Lange, A., and Schulden, P. (2020). The role of the CFO of an industrial company: an analysis of the impact of blockchain technology. *Future Internet* 12, 1–16. doi: 10.3390/fi12080128
- Shah, D. (2021). By the numbers: MOOCs in 2021—class central. Available at: <https://www.classcentral.com/report/mooc-stats-2021/> (Accessed: June 11, 2021)
- Shah, C., Davtyan, K., and Nasrallah, I. (2023). Artificial intelligence-powered clinical decision support and simulation platform for radiology trainee education. *J. Digit. Imaging* 36, 11–16. doi: 10.1007/s10278-022-00713-9
- Shorey, S., Ang, E., Yap, J., Ng, E. D., Lau, S. T., and Chui, C. K. (2019). A virtual counseling application using artificial intelligence for communication skills training in nursing education: development study. *J. Med. Internet Res.* 21:e14658. doi: 10.2196/14658
- SIC (2021). SIC online Ideathon. (Accessed January 25, 2023) Available at: <https://s-ic.cz/en/sic-online-ideathon-welcome-office/>
- Siemens, G. (2012). Learning analytics: envisioning a research discipline and a domain of practice. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. 4–8.
- Steinmetz, S. (2022). How to Master Digital Facilitation: Best practices and an easy guide for online workshops. *SAP Community - Personal Insights*. <https://blogs.sap.com/2022/11/09/how-to-master-digital-facilitation-best-practices-and-an-easy-guide-for-online-workshops/>
- Suárez-Brito, P., López-Caudana, E. O., Baena-Rojas, J. J., and Ramírez-Montoya, M. S. (2022). Eliciting complex thinking through open educational resource projects. *J. Soc. Stud. Educ. Res.* 13, 56–77.
- Telnov, V., and Korovin, Y. (2019). Semantic web and knowledge graphs as an educational technology of personnel training for nuclear power engineering. *Nucl. Energy Technol.* 5, 273–280. doi: 10.3897/nucet.5.39226
- Tsalapatas, H., Heidmann, O., Pata, K., De Carvalho, C. V., Bauters, M., Papadopoulos, S., et al. (2019). Teaching design thinking through gamified learning. *CSEDU 2019—Proceedings of the 11th International Conference on Computer Supported Education*, 278–283
- Tubaro, P. (2022). Learners in the loop: hidden human skills in machine intelligence. *Sociol. Lav.* 163, 110–129. doi: 10.3280/SL2022-163006
- Uesugi, M., and Moriyama, T. (2020). Regional disaster prevention with cloud-based GIS mapping and location-based gaming: disaster prevention go!. In *Proceedings of the 22nd IAHR-APD Congress 2020*.
- United Nations (2015) Transforming our world by 2030: A new agenda for global action zero. *Draft of the outcome document for the UN summit to adopt the post-2015 development agenda*. New York: United Nations. Available at: <https://sdgs.un.org/2030agenda>
- University of Washington (2016). C21: the center for 21st century liberal learning. *Ideathon*. (Accessed 25 January, 2023) Available at: <https://c21.washington.edu/our-programs/ideathon>
- University of Waterloo (n.d.). *University relations: Determining event objectives*. <https://uwaterloo.ca/university-relations/resources/campus-eventplanning/planning-considerations>
- Vázquez-Parra, J. C., Cruz-Sandoval, M., and Carlos-Arroyo, M. (2022). Social entrepreneurship and complex thinking: a bibliometric study. *Sustainability* 14:13187. doi: 10.3390/su142013187
- Wang, Y. B., and Ho, C. W. (2017). No money? No problem! The value of sustainability: social capital drives the relationship among customer identification and citizenship behavior in sharing economy. *Sustainability* 9:1400. doi: 10.3390/su9081400
- World Economic Forum (2015). *Understanding the sharing economy*. Retrieved from: https://www3.weforum.org/docs/WEF_Understanding_the_Sharing_Economy_report_2016.pdf
- Xia, Q., Chiu, T. K., Zhou, X., Chai, C. S., and Cheng, M. (2022). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Comput. Educ.: Artif. Intell.* 4:100118. doi: 10.1016/j.caeai.2022.100118
- Yang, C., Huan, S., and Yang, Y. (2020). A practical teaching mode for colleges supported by artificial intelligence. *Int. J. Emerg. Technol. Learn.* 15, 195–206. doi: 10.3991/ijet.v15i17.16737

- Yang, D., Zhou, J., Shi, D., Pan, Q., Wang, D., Chen, X., et al. (2022). Research status, hotspots, and evolutionary trends of global digital education via knowledge graph analysis. *Sustainability* 14:15157. doi: 10.3390/su142215157
- Yu, J. (2021). Academic performance prediction method of online education using random Forest algorithm and artificial intelligence methods. *Int. J. Emerg. Technol. Learn.* 16, 45–57. doi: 10.3991/ijet.v16i05.20297
- Yudina, V., Skrupskaya, Y., Taratukhin, V., Kozlova, E., and Pulyavina, N. (2021). Bringing together engineering and management students for a project-based global idea-thon: towards next-gen design thinking methodology. Paper Presented at 2021 ASEE Virtual Annual Conference Content Access, Virtual Conference
- Zakharkina, L., Rubanov, P., Alibekova, B., Zakharkin, O., and Moldashbayeva, L. (2022). The impact of digital transformation in the accounting system of fuel and energy complex enterprises (international experience). *Int. J. Energy Econ. Policy* 12, 152–161. doi: 10.32479/ijeep.13317
- Zhao, X. X., Wu, S. P., Wang, J. Y., Gong, X. Y., He, X. R., Xi, M. J., et al. (2020). Comparison of multiple quantitative evaluation indices of theoretical knowledge and clinical practice skills and training of medical interns in cardiovascular imaging using blended teaching and the case resource network platform (CRNP). *Med. Sci. Monit.* 26:e923836. doi: 10.12659/MSM.923836
- Zhu, Q., Zhang, F., Huang, Y., Xiao, H., Zhao, L. Y., Zhang, X. C., et al. (2022). An all-round AI-chemist with a scientific mind. *Natl. Sci. Rev.* 9:nwac190. doi: 10.1093/nsr/nwac190



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Maila Pentucci,
University of Studies G. d'Annunzio Chieti and
Pescara, Italy
Lina Kaminskienė,
Vytautas Magnus University, Lithuania

*CORRESPONDENCE

Claudia Camacho-Zuñiga
✉ claudia.camacho@tec.mx

RECEIVED 31 March 2023

ACCEPTED 25 May 2023

PUBLISHED 15 June 2023

CITATION

Camacho-Zuñiga C, Peña-Becerril M,
Cuevas-Cancino MO and
Avilés-Rabanales EG (2023) Gains from the
transition from face-to-face to digital/online
modality to improve education in emergency.
Front. Educ. 8:1197396.
doi: 10.3389/feduc.2023.1197396

COPYRIGHT

© 2023 Camacho-Zuñiga, Peña-Becerril,
Cuevas-Cancino and Avilés-Rabanales. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Gains from the transition from face-to-face to digital/online modality to improve education in emergency

Claudia Camacho-Zuñiga^{1,2*}, Maritza Peña-Becerril²,
Maria de la O Cuevas-Cancino² and
Eréndira Gabriela Avilés-Rabanales²

¹Institute for the Future of Education, Tecnológico de Monterrey, Monterrey, Mexico, ²School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico

Education during emergencies demands attention from both institutions and governments, as the incidence of disasters and crises continues to increase worldwide due to various challenges that threaten humanity, including climate change, terrorism, food and water scarcity, energy crises, poverty, and inequality. Emergency remote teaching (ERT) is a critical factor in providing academic opportunities, building resilience, and promoting recovery for affected individuals. This research examines the challenges associated with transitioning from face-to-face to online/digital lessons, according to the perceptions of 105 professors in a Mexican private university, 2 months after mandatory online/digital teaching was implemented. By analyzing data collected through a survey using statistical and data mining techniques, we identified factors that should be considered when designing future education in emergency situations. The study found that while most professors positively evaluated their performance in ERT, they still preferred face-to-face teaching due to the lack of personal interaction and distractions. Professors adapted to ERT with technological training provided by the institution, but reported an increase in workload and the importance of the balance between job responsibilities and family issues. The study also revealed that professors' performance in ERT depended on the conditions of their teaching space at home. Furthermore, the study suggested the need for the development of an improved assessment framework for online evaluation methods, and integration of the most effective online and face-to-face teaching strategies for everyday academic contexts. The sudden change in education modality represented a fantastic opportunity to introduce teaching-learning-evaluating strategies and technologies while forcing institutions to evolve into the digital environment. Therefore, every education stakeholder should identify and implement best practices of the online and face-to-face modalities to improve future education in emergencies.

KEYWORDS

faculty perceptions, professors' performance, digital education, educational innovation, higher education, remote education in pandemic

1. Introduction

The incidence of disasters and emergencies is increasing globally, ranging from pandemics to wars. These challenges pose a significant threat to humanity, including climate change, terrorism, food and water scarcity, energy crises, poverty, and inequality (United Nations, 2022). The rapid advancement of technology has highlighted the relevance of acquiring knowledge and its practical application as a critical means of addressing these challenges, mitigating disasters, and reducing their impact. In light of this, providing education during emergency situations is of utmost importance for both present and future generations.

Delivering education in emergency situations is crucial for providing educational opportunities and support to those affected by crises. In natural disasters, conflicts, epidemics, and other difficulties, education helps students by providing a sense of stability and assists them in mitigating the long-term effects of the emergency. Furthermore, education can also foster resilience and promote recovery for the future (Torani et al., 2019). It is noteworthy, however, that the success of education relies not only on institutions but primarily in educators, who are themselves susceptible to the trauma and stress associated with emergencies.

Amid the COVID-19 pandemic, lockdown represented a unique challenge to educators, as they were the only group of professionals who had no other option but to work remotely (Allen et al., 2020). Many employees in other sectors had the option of not leaving their working places or returning to their workplaces under safety protocols. However, professors, teachers, and instructors did not have this option since most governments around the Globe implemented social distance and closing of all educational facilities (UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, Economic Commission for Latin America and the Caribbean, 2020). In March 2020, Mexico announced the suspension of all academic activities and closure of schools in all levels. This decision suddenly replaced face-to-face education with distance education through several formats and platforms, with or without the use of technology (UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, Economic Commission for Latin America and the Caribbean, 2020; Area and Adell, 2021).

This study investigated professors' perception 2 months after the urgent implementation of online classes. The research question that guided this analysis was: how did faculty perceive their academic performance during the transition of their teaching to digital modality? The first purpose was to disseminate among the teaching community how other colleagues experienced this transition. This would help identify common challenges, learn from our peers, and create a sense of community among faculty members. The second purpose was to allow educational institutions, decision-makers, and policymakers to identify the best strategies to support faculty and ensure not only continuity during emergencies but also academic quality in any circumstances. The second purpose of the study was to identify the best strategies to support faculty and ensure academic quality in any circumstances. This is particularly important, as future emergencies could include a range of crises such as wars, climate change consequences, water, and food shortages, among others. In fact, according to Duke Global Health Institute (Penn, 2021), large pandemics are even more likely to happen than previously thought.

2. Theoretical framework

Due to the pandemic that affected the Globe in 2020, schools, colleges, and universities had to close their doors. As a response, this triggered the deployment of distance education modalities through a variety of formats and platforms (UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, Economic Commission for Latin America and the Caribbean, 2020). Some educational institutions were able to quickly transfer their face-to-face classes to online or digital classes despite having little time to prepare (Huang et al., 2020). In many other educational institutions in Latin America, due to the digital divide or the lack of economic and technological resources (Ferri et al., 2020) some other distance education modality was implemented, e.g., through open television broadcast. In worst-case scenarios, some institutions stopped providing school services and attention to students all together (Torrás Virgili, 2021).

Educational institutions faced a crisis in 2020 as communities and governments continued to make difficult decisions about education during COVID-19 pandemic (Allen et al., 2020; Flack et al., 2020). Examples of these decisions include the use of video conferencing for ensuring the delivery of education including lecturing, learning activities, and assessment; the use of video conferencing for faculty collaboration and academic collaborative work; the fostering of learning management systems (LMS) to create learning environments for the tracking of academic performance of students (Bond et al., 2021); the creation of web repositories for digital learning materials; the development and proliferation of apps and software to facilitate the creation of videos, infographics, animated videos, podcast, and other multimedia content; the production and distribution of TV programs with learning content (nationally and locally); reduced-price agreements with telecommunications and digital companies to provide their services for educational purposes; and the organization of conferences, academic and training activities as webinars, and other online activities (Area and Adell, 2021).

The educational environment changed radically; it was forced to reinvent its learning, teaching, and evaluation processes. According to Miguel Román (2020) and Alcántara Santuario (2020), the shift to online education was a difficult and drastic adjustment for educational institutions. This global crisis triggered a rethinking of educational service delivery at all levels (Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe, 2020). The intensive use of all kinds of platforms and resources technology to ensure continuity of learning is the boldest experimenting the field of educational technology, although unexpected and unplanned.

To learn from this crisis, while overcoming it, it was necessary to analyze the challenges involved, especially for online education. Topics like the attention to students, the use of educational technologies to support the teaching process and the well-being and needs of teachers and professors, needed to be analyzed. According to Sánchez Mendiola et al. (2020) teachers left the traditional classroom, to which they have been accustomed for decades, to become forced users of the technological tools that exist to interact at a distance between themselves and their students, while dealing with the pressures of confinement and its economic, health, and emotional consequences. In this address, See et al. (2020) affirmed that the workload of teachers is directly related to their well-being, besides; Hayes et al. (2020) reported that working from home can cause greater stress.

For Miguel Román (2020), successfully dealing with this educational crisis implied that both teachers and students develop self-learning, autonomy, and socio-emotional skills. The challenge was not easy, so it is important to recognize that many teachers and professors, along with their institutions and students, found ways to innovate. The foregoing led education to a renewal; Flack et al. (2020) recognized and celebrated the great work that all stakeholders in education did. However, every innovation in education must be analyzed and validated. We must provide strong evidence of true improvement in each innovation, its circumstances and even identify their limitations; or, conversely, identify those innovations that create more problems than solutions (Moreno, 2021).

The definition of Emergency Remote Teaching (ERT) clearly states that it is a method of teaching meant to be used only in emergency situations. Due to the magnitude of the COVID-19 Pandemic, we do not have a historical point of comparison to either the short term or long-term impacts of this online methodology on both students and professors, therefore, the need to have as much information on these impacts becomes clear (Ezra et al., 2021; Stewart, 2021).

In fact, multiple studies have been published concerning the impact on both the teaching and learning process (Allen et al., 2020; Ferri et al., 2020; Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe, 2020), as well as on the development, health, and well-being of students (Area and Adell, 2021; Camacho-Zuñiga et al., 2021). However, it was important to find out how well-prepared every educational stakeholder was for the online modality. Concerning teachers and professors, it is recommended to train them in ERT (Trust and Whalen, 2020), especially after identifying several areas of opportunity amid education during COVID-19 lockdown (Aicántara Santuario, 2020; UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, Economic Commission for Latin America and the Caribbean, 2020). This makes our analysis relevant, since it specifically focuses on professors' perspective—on how they perceived their online teaching performance, focusing attention into opportunities for learning, innovation, and improvements in higher education during an emergency.

3. Methodology

3.1. Data collection

The private Mexican university studied here had successful experiences in online instruction before the COVID-19 lockdown, such as distance programs with synchronous and asynchronous courses. These courses were implemented in Mexico City after the earthquake in 2017. This allowed the training of 14,000 professors in a fast, pertinent, and concrete way, in relation to the use of platforms such as Zoom or Microsoft Teams and the fundamental strategies for online teaching. The university began remote online instruction on Monday, March 23, 2020, providing technical support and monitoring to the professors.

The faculty perspective on the mandatory transition from face-to-face modality to online modality after the COVID-19 breakout in

Mexico was analyzed through a sample of 105 professors. Faculty from the Schools of Humanities, Engineering, Business, Architecture and Social Sciences and Government were invited via institutional email to answer a survey between April and May 2020. They were informed of the purpose of the study, their optional participation, and the confidentiality of their answers. Participants' adscription corresponded to Campi in the cities of Toluca, Puebla, and Mexico City in Mexico (all located in the south-center of Mexico). Their age range was 25–64 years with 1–43 years of teaching experience. Among the participants, 61 identified as males and 44 as females and all of them held a masters and/or doctorate degree.

The participants of this research experienced online modality during the COVID-19 emergency, as a teaching and learning situation in which undergraduates were in a distance learning situation and not face-to-face with their professors. Undergraduates and professors interacted synchronously through videoconferencing technology and professors used a LMS as an aid for planning, informing, sharing, and evaluating learning.

3.2. Survey

We collected data through an online survey, powered by Google Forms, which included 10 closed-ended and four open-ended questions written in Spanish. At the beginning of this survey, participants informed about their degrees, years of teaching experience, and general socio/demographic aspects. The core questions focused on their perspective of the transition from the face-to-face modality to the online modality. The closed-ended questions under analysis in this research are shown in Table 1. The open-ended questions under analysis are: (9) Taking into consideration your teaching activities, what are the fundamental differences that you find between the two teaching experiences: face-to-face lectures and online lectures? (10) Regarding students' performance, what is the main difference between face-to-face lectures and online lectures? (13) What is the main advantage in online lectures? (14) What is the main disadvantage in online lectures? Questions 4, 11, and 16 were topics of a different publication.

3.3. Analysis

Data gathered from closed-ended questions were analyzed using descriptive and inferential statistics as well as data mining (*k*-means clustering); while that from open-ended questions was analyzed using text mining. Spearman's rank correlation coefficient (Null hypothesis: there exists no correlation between variables, Alternative hypothesis: there exists correlation between variables, level of significance 0.05) was evaluated for different variables assessment of the online modality (Tables 2–4) using Minitab Ver 19.2020.2.0.

K-means clustering is a machine unsupervised learning tool used to find relevant information beneath the data; it provides an insight of data that needs to be explained by the experimenter (Essa and Mojarad, 2022). Its implementation required transforming categorical data into numerical data (0–4). It was executed with 11 features (degree, online teaching evaluation, perceived change in workload, time to adapt to the online modality, perceived online teaching

TABLE 1 Closed-ended questions for faculty participants about the mandatory transition from face-to-face modality to online modality after COVID-19 lockdown in Mexico.

Question	Options				
1. At home, how is your workspace to deliver online lessons? Is it isolated, illuminated and free from distractions and noise?	0. I do not have a specific space for my lecture	1	2	3	4. It is a bright, isolated and distraction-free space
2. Taking into consideration your experience in the online modality, how do you evaluate the following aspects? (See Figure 1)	Very bad	Bad		Good	Excellent
3. Taking into consideration your previous answers, in which of the two modalities do you perceive the following characteristics are better? (See Figure 2)	Definitely face-to-face modality	A little better in face-to-face modality	Equivalent in both	A little better in the online modality	Definitely online modality
5. Taking into consideration a subject that you deliver regularly in presence, for each hour of lecture, how much time did you dedicate to your preparation?	Less than 1 h	Between 1 and 2 h	Between 2 and 3 h	More than 3 h	I have no arguments to answer
6. Taking into consideration a subject that you regularly deliver in face-to-face, and that you are delivering in the online modality, for each hour of lecturing, how much time do you dedicate now to its preparation?	Less than 1 h	Between 1 and 2 h	Between 2 and 3 h	More than 3 h	I have no arguments to answer
7. How quickly did you adapt to the change of face-to-face modality to online modality?	I have not adapted	Slowly		Moderate-ly Fast	I adapted quickly
8. Which type of lecturing do you like the most: Face-to-face modality or online modality?	Definitely face-to-face modality	A little more face-to-face modality	It is indistinct to me	A little more online modality	Definitely online modality
12. Compared to a face-to-face lecture, how has your workload changed?	It decreased a lot	It decreased a little	It has stayed the same	It increased a little	It increased a lot
15. When your online lessons end, you say goodbye to your alumni and close your Zoom, how satisfied do you feel with your teaching?	0 Not satisfied at all	1	2	3	4 Extremely satisfied
17. How much harmony is there between your family issues and your job responsibilities working from home?	0 Very little	1	2	3	4 A lot

Numbering reflects actual order in the survey.

performance, school, harmony between job and family issues, teaching experience, satisfaction of their own online teaching performance, working space quality, and gender) without dimensionality reduction. Finally, its performance was evaluated through silhouette coefficient.

Answers to the open-ended questions were cleaned (correct spelling; removing accent mark, stop words and plurals; unifying synonyms; and replacing verbs by their infinitive) prior to the text analysis performed using Voyant Tools (Sinclair and Rockwell, 2016).

4. Results

From an overall perspective, after 5–6 weeks of mandatory closure of face-to-face instruction, faculty evaluated the online modality positively. Figure 1 reports the frequency of answers to assess the characteristics of the online modality through a heatmap (higher frequency corresponds to higher intensity and vice versa). It is noteworthy that “Good” got the highest frequency for every analyzed characteristic of the online teaching modality, which evidences a

	very bad	bad	good	excellent
Your own performance	0	6	70	29
Undergraduates' participation	3	16	72	14
Undergraduates' fulfillment in assignments	2	3	61	39
Integrity during evaluation	4	21	64	16
Undergraduates' learning	1	13	72	19
Quality in undergraduates' contribution	0	24	57	24
Undergraduates' punctuality and attendance	2	6	57	40
Fluidity in sessions	0	3	66	36
Professor-undergraduate interaction	3	13	54	35
Interaction between undergraduates	5	29	55	16
Syllabus and contents coverage	1	8	48	48
Undergraduates' attitudes towards learning	1	16	66	22
Assessment effectivity	2	18	69	16
Innovation in lectures	1	15	59	30

FIGURE 1

Frequency in professors' responses ($n=105$) to the question: "Taking into consideration your experience in the online modality, how do you evaluate the following aspects?" after mandatory shift from face-to-face instruction to online/digital instruction during COVID-19 quarantine. Higher intensity corresponds to higher frequency.

growth in professors' competencies far beyond the sudden shift from face-to-face modality during the COVID-19 emergency and the challenges that came with it.

Concerning the attribute of "excellent," professors mentioned coverage of the syllabus, undergraduate's punctuality and attendance, fulfillment in assignments, and fluidity of the sessions in the online modality with the highest frequency. Participants evaluated syllabus and contents coverage as excellent, which corresponds with the perception of fluidity during the lessons and the fulfillment of assignments. These results might be related to the multiple tools and sources of learning available for the undergraduates in the World Wide Web, especially during the global pandemic; learners can review and try out different didactic sources, such as class recordings, blogs, videos, papers, textbooks, notes, etc. On the other hand, participants evaluated punctuality of their alumni as excellent, which reflected the ease of access to the classroom in its virtual mode.

Under professors' perception, the most mentioned challenges of online/digital teaching were identified as the interaction between undergraduates, quality in their contribution/participation, integrity during evaluation, and assessment effectiveness. The first two were related to social interactions in the online modality; these aspects were fundamental for the teaching and learning process since social relevance has been highlighted as an intrinsic motivation to learn (Ryan and Deci, 2000; Savelsbergh et al., 2016). The last two challenges were related to assessment and evaluation, which were rated as "bad" under the participants' perspective. This result has several implications, which will be discussed later.

After comparing the teaching-learning process, respondents assessed face-to-face modality equivalents with online/digital modality in eight out of 14 analyzed characteristics. Figure 2 shows a heatmap of the frequency in professors' response to the question: "Considering your previous answers, in which of the two modalities

do you perceive the following characteristics are better?" after mandatory interruption of face-to-face instruction during the COVID-19 quarantine. Participants evaluated both modalities' equivalents in undergraduates' fulfillment in assignments, achievement of learning content, undergraduates' punctuality and attendance and fluidity during sessions. They preferred the face-to-face modality over the online/digital modality in four factors; in decreasing order of frequency, they mentioned: (1) interaction between undergraduates, (2) undergraduates' participation, (3) integrity during evaluation, and (4) personal interaction with undergraduates. In contrast, online modality was definitely preferred over face-to-face modality due to the innovation involved in lectures. In the answers to Question 8, "Which type of lecturing do you like the most: Face-to-face modality or online modality?," 21.9% of the participants preferred online lectures over face-to-face lectures, 19.0% of the participants had no preference for any of the modalities, while 59.0% of the participants liked the face-to-face modality better than the online modality. This discrepancy between specific characteristics and overall evaluation might be a consequence of those facts and issues that fall outside of the analysis.

How professors assessed their own performance during the start of the COVID-19 emergency in online modality was correlated with most of the variables analyzed herein, but higher correlations were found to relate to undergraduates' participation and the quality of their contributions. Table 2 reports the Spearman rank correlation analysis of the professors' performance and the different characteristics of the online modality, as assessed in the participants' answers to the question: "Taking into consideration your experience in the online modality, how do you evaluate the following aspects?" It can be noted that undergraduates' punctuality and attendance is the only variable that did not correlate with the professors' assessment of their own performance. On the contrary, professors' performance in the online

	Definitely face-to-face modality	A little better in face-to-face modality	Equivalent in both	A little better in the online modality	Definitely online modality
Your own performance	22	33	36	10	4
Undergraduates' participation	38	40	18	7	2
Undergraduates' fulfillment in assignments	14	19	58	11	3
Integrity during evaluation	35	31	38	1	0
Undergraduates' learning	13	44	41	6	1
Quality in undergraduates' contribution	18	32	43	10	2
Undergraduates' punctuality and attendance	17	15	46	15	12
Fluidity in sessions	19	25	45	11	5
Professor-undergraduate interaction	35	32	25	8	5
Interaction between undergraduates	46	31	22	6	0
Achievement of learning content	15	29	55	5	1
Undergraduates' attitudes towards learning	17	38	43	7	0
Assessment effectivity	22	42	34	5	2
Innovation in lectures	11	12	31	33	18

FIGURE 2

Frequency in professors' response to the question: "Considering your previous answers, in which of the two modalities do you perceive the following characteristics are better?" after mandatory closure of face-to-face instruction during COVID-19 quarantine. Higher intensity corresponds to higher frequency.

TABLE 2 Spearman's rank correlation (Significance level: 0.05) coefficient between professors' perception of their own performance (Question 2A) and the rest of perceptions about the online modality, according to Question 2: "Taking into consideration your experience in the online modality, how do you evaluate the following aspects?"

Professors' performance versus:	Correlation	p value	Conclusion
Undergraduates' participation	0.432	0.000	Correlated
Undergraduates' fulfillment in assignments	0.364	0.000	Correlated
Integrity during evaluation	0.376	0.000	Correlated
Undergraduates' learning	0.422	0.000	Correlated
Quality in undergraduates' contribution	0.424	0.000	Correlated
Undergraduates' punctuality and attendance	0.029	0.728	---
Fluidity in sessions	0.418	0.000	Correlated
Personal interaction with undergraduates	0.289	0.003	Correlated
Interaction between undergraduates	0.265	0.007	Correlated
Achievement of learning content	0.379	0.000	Correlated
Undergraduates' attitudes toward learning	0.276	0.005	Correlated
Assessment effectivity	0.380	0.000	Correlated
Innovation in lectures	0.401	0.000	Correlated

modality showed the highest correlation with undergraduates' participation, quality in undergraduates' contribution during lectures and undergraduates' learning. The positive correlation with these variables reflects how the professors' attention was student-centered, and therefore it becomes clear that personal interaction between professors and undergraduates is fundamental for the teaching-learning process.

According to the perspective of the professors who participated in this study, there was a correlation between their teaching performance and the characteristics of the workspace at home and the harmony with their families. Table 3 reports the Spearman correlation analysis

between professors' perception of their own performance (Question 2) and answers given to Questions 1, 15, and 17. It is valuable to note that characteristics of the space at home dedicated to online teaching as well as family harmony with work correlate positively with instructor's perception of their performance and are relevant when considering future education scenarios during emergencies.

Professors' satisfaction with their online teaching correlated with the harmony they perceived between their families, their job, and with having an appropriate workspace at home. Table 4 shows the clustering K-means results performed without dimensionality reduction on 11 features, where three of them accounted for 58% of

TABLE 3 Spearman's rank correlation coefficients between professors' perception of their own performance (Question 2A) and characteristics of the workspace at home, personal satisfaction about teaching, and work-family harmony.

Professors' performance (Question 2) versus:	Correlation	p value	Conclusion
1. At home, how is your workspace to deliver online lessons? Is it isolated, illuminated and free from distractions and noise?	0.254	0.009	Correlated
15. When your online lessons end, you say goodbye to your alumni and close your Zoom, how satisfied do you feel with your teaching?	0.530	0.000	Correlated
17. How much harmony is there between your family issues and your job responsibilities working from home?	0.208	0.034	Correlated

Significance level: 0.05.

TABLE 4 Clustering K-means results (11 features, without dimensionality reduction, and Silhouette coefficient of 0.1929).

	Professors in cluster	Family-job harmony means		Working space quality means		Online-performance satisfaction means	
		Cluster	Population	Cluster	Population	Cluster	Population
Cluster 1 Low online teaching performance	21 (20%)	2.1	3.23	1.76	3.1	2.57	3.26
Cluster 2 High online teaching performance	84 (80%)	3.51		3.42		3.46	

Participants in Cluster 1 perceived their online-teaching performance on average 21% lower than the population mean; while participants in Cluster 2 perceived their online-teaching performance on average 5% higher than the population mean.

the variance: harmony between job and family issues (23%), quality of the working space at home (20%), and satisfaction with their own online teaching performance (15%). Classification resulted in two clusters with a silhouette coefficient of 0.1929. Cluster 1 included 20% of the participants showing a lower quality workspace at home and scarce harmony between family issues and their job, which resulted in a lower perception of their online teaching performance. Cluster 2 included 80% of the participants showing a higher quality workspace and more harmony between family and work, which concurs with a higher perception of their online teaching performance. These results point out professors' necessity for a space and time especially dedicated to their academic activities, far from distractions, and family concerns. This might provide evidence for decision-makers to return every academic activity inside institutional facilities, whenever possible.

While none of the participants reported not having adapted to the online modality, professors' time to adjust correlated positively with their perception of their teaching performance and achievement of the learning content. Table 5 reports the Spearman correlation analysis of the professors' perception of the time needed to adapt to the online modality (Question 7) and their perception of the different aspects evaluated in Question 2 about this modality. 61.9% of the participants considered that they quickly adapted to the online modality; 33.3% of them adapted moderately fast and only 4.76% of them adapted slowly to this modality. Among the highest positive correlations with the time needed to become accustomed to the online modality, two are related with the professor: the achievement of learning content and their own performance. The other three characteristics are related with the undergraduates: their participation, their fulfillment in assignments, and the quality of their contribution during lectures.

In agreement with Hayes et al. (2020), professors who participated in this research perceived that the workload related to preparing lessons and evaluating increased when changing from face-to-face to

online instruction due to the COVID-19 lockdown. Their answers to Question 12 (Table 1) reflected that for 7.6% of the participants the workload decreased in comparison to face-to-face lectures, for 7.6% of the participants the workload remained the same, but for 84.8% of the participants the workload increased. These results can be understood considering that participants' expertise focused on face-to-face modality, and they identified the online modality as highly innovative. Besides, the survey time span was too close to the beginning of the COVID-19 lockdown to allow professors to go through a natural learning curve and adapt to the new requirements of online/digital teaching, as has been reported in the literature.

Participants perceived interaction as the main difference between face-to-face and online modalities. Figure 3 shows the word cloud of the top 20 words in the answers to Question 9. The term "interaction" was the most significant term in this word cloud. Table 6 reports some relevant terms and their highest correlations to other terms present in the open-ended questions. Results suggest that participants identified the interaction between lecturer-students and student-student as being relevant and must be incorporated in lectures regardless of the modality. In agreement with previous results, the participants identified the workload as a difference between modalities and there was a correlation with the online modality.

Participants identified "attention" as a difference in undergraduates' performance in the face-to-face and online modalities. Figure 4 shows the word cloud of the top 20 words in the answers to Question 10 with the most significant term being "attention," which is strongly correlated with "environment"; this might be linked to the difficulty students had to focus their attention during the online lessons due to the multiple sources of distractions in their homes and in the WWW.

Professors who participated in this research mentioned less transportation as the main advantage of the online modality. Figure 5 shows the word cloud of the top 20 terms present in the

TABLE 5 Spearman's rank correlation coefficient between professors' perception of the time to adapt to the shift from face-to-face modality to online modality (Question 7) and their perception of the online modality (Question 2).

Time to adapt (Question 7) versus:	Correlation	p value	Conclusion
Professors' performance	0.360	0.000	Correlated
Undergraduates' participation	0.307	0.002	Correlated
Undergraduates' fulfillment in assignments	0.303	0.002	Correlated
Integrity during evaluation	0.150	0.129	---
Undergraduates' learning	0.280	0.004	Correlated
Quality in undergraduates' contribution	0.287	0.003	Correlated
Undergraduates' punctuality and attendance	0.146	0.140	---
Fluidity in sessions	0.260	0.008	Correlated
Personal interaction with undergraduates	0.126	0.203	---
Interaction between undergraduates	0.073	0.464	---
Achievement of learning content	0.287	0.003	Correlated
Undergraduates' attitudes toward learning	0.265	0.007	Correlated
Assessment effectivity	0.151	0.126	---
Innovation in lectures	0.279	0.004	Correlated

Significance level: 0.05.

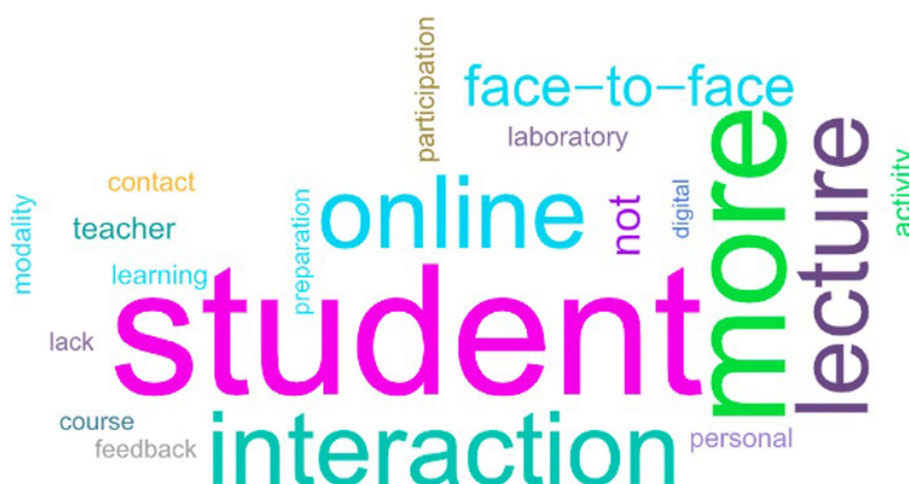


FIGURE 3

Top-20 terms in the answers of the participants ($n=105$) to the open-ended question: "Taking into consideration your teaching activities, what are the fundamental differences that you find between the two teaching experiences: face-to-face lecture and the online lecture?" (Question 9).

answers to Question 9, concerning the main advantage of the online modality. The results identified that the most significant terms were "home," "transportation," "family," and "recorded." The correlations shown in Table 5 suggest that less transportation during the online modality led to more interaction with their families and that the possibility to record the lessons facilitated comprehension in the undergraduate students.

According to the perceptions of the participants, the lack of interaction with the undergraduates and the absence of laboratories constituted the disadvantages of the online modality. Figure 6 shows the word cloud of the top 20 terms in the professors' answers to Question 14. Table 6 shows that "interaction" is clearly related to "lack" and "student." It can be concluded that the online modality strongly limited the instructor-student and student-student

interactions, which agrees with the previously discussed answers to Question 2 and 9. The other significant term that appeared in this word cloud is "laboratory." In higher education programs, laboratories play an important role in the development of specific competencies, with several of them requiring experimentation and use of materials and equipment that are not always available in the virtual environment.

5. Discussion

Worldwide, the lockdown to diminish the spread of COVID-19 strongly affected the way professors delivered their teaching (Allen et al., 2020). In March 2020, Mexican higher education institutions promoted lockdown as a measure to diminish the spread of

TABLE 6 Pairs of terms with highest correlation in the professors' open-ended answers to the fundamental differences between face-to-face and online modalities (Question 9), the main differences between the undergraduates' performance (Question 10), as well as the advantages (Question 13) and disadvantages of the online modality (Question 14).

Question	Term 1	Term 2	Correlation
Question 9	Important/importancia	Interaction/interacción	0.7170
	Incorporation/incorporación	Interaction/interacción	0.7170
	Student/estudiante	Interaction/interacción	0.6949
	Advance/avance	Interaction/interacción	0.6411
	Workload/carga	Online/en línea	0.6201
	Contents/contenido	Online/en línea	0.6201
	Communication/comunicación	Face-to-face/presencial	0.6343
	Camera/cámara	Evaluation/evaluación	0.6124
	Complicated/complicado	Evaluation/evaluación	0.6124
	Contents/contenidos	Evaluation/evaluación	0.6124
	Coexistence/cohabitar	Evaluation/evaluación	0.6124
Question 10	Trust/confianza	Performance/desempeño	0.7638
	Environment/ambiente	Attention/atencion	0.7071
Question 13	None/ningun	Transportation/traslado	0.8023
	Less/menos	Transportation/traslado	0.7981
	Learning/aprendizaje	Transportation/traslado	0.7158
	To eat/comer	Transportation/traslado	0.7158
	Knowledge/conocimiento	Transportation/traslado	0.7158
	Understanding/entendimiento	Recorded/grabada	0.6784
Question 14	Lack/falta	Interaction/interacción	0.7673
	Student/estudiante	Interaction/interacción	0.6710

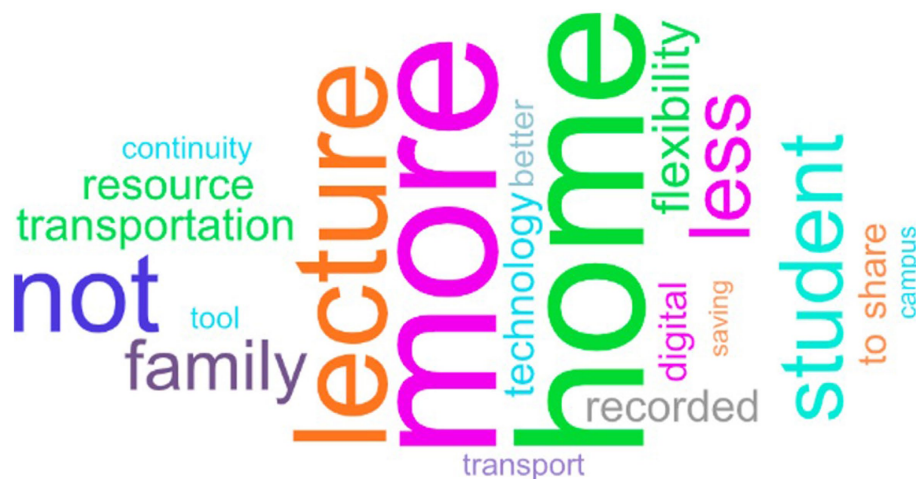


FIGURE 4

Top-20 terms in the answers of the participants ($n=105$) to the open-ended question: "Regarding students' performance, what is the main difference between face-to-face lectures and online lectures" (Question 10).

COVID-19. Mandatory closure of all educational facilities and schools forced every teaching activity from face-to-face modality into an online modality. Despite their previous expertise, professors were forced to adapt their face-to-face teaching methods to the technological tools and requirements of the online modality in a

two-week timespan. Besides that, professors had to deal with the use of digital tools and platforms as well as having to teach from home (Trust and Whalen, 2020).

From an overall perspective, after a month of mandatory closure of face-to-face instruction, faculty from this private Mexican

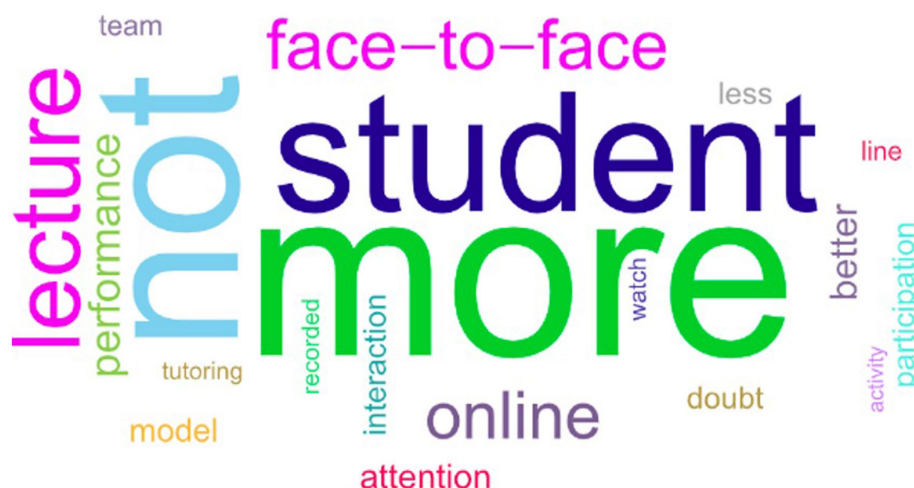


FIGURE 5

Top-20 terms in the answers of the participants ($n=105$) to the open-ended question: "What is the main advantage in online lectures?" (Question 13).



FIGURE 6

Top-20 terms in the answers of the participants ($n=105$) to the open-ended question: "What is the main disadvantage in online lectures?" (Question 14).

university evaluated the online modality positively. This might have reflected the previous experience of the institution with online learning, distance programs, and in emergencies such as the damage caused to the facilities after an earthquake in 2017. However, during the global COVID-19 emergency, the university trained its professors in the use of technologies and strategies for the online modality and launched it within 1 week. This result also reflected on the micro-environment inside of each class, as a continuously adapting organism where the professor designed the best suitable track under his/her academic freedom.

From the perspective of professors used to a face-to-face modality, some of the challenges involved in the online modality were the interaction with the undergraduate students (Ferri et al., 2020), the laboratory instruction, the attention of undergraduates, and the efficiency of evaluation (UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, Economic Commission for Latin America and the Caribbean, 2020). The lack of

social interaction and communication was also recognized by students and reported in literature (Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe, 2020). Student-student interaction promotes significant learning; while instructor-student interaction is relevant to tracking the learning process, and even facilitates that the undergraduates focus their attention. Therefore, personal interaction must be favored during lessons through any tool, e.g., synchronic sessions, collaborative activities, and breakout rooms. Regarding the evaluation efficiency, professors must be aware of the multiple sources of information available for the students not only during professional life, but also while taking exams. Instead of mimicking traditional evaluation into a digital environment, a new and improved assessment framework needs to be developed. Evaluation in online courses is a relevant topic that requires further research.

Participants positively evaluated coverage of the syllabus and fluidity during lessons in the online modality; in addition to

recognizing that recordings of synchronic lessons were helpful to promote students' understanding. This might reflect the multiple digital tools at hand in online learning, e.g., videos, digital libraries, presentations, simulators, and graphing calculators; unfortunately, it might also be related to the decreased or null interaction with students, which led to high fluidity in professors' lecturing.

Over 95% of the faculty that used to be in a face-to-face or in-person modality adapted fast or moderately fast to the requirements of this new modality; however, after 2 months of online instructions, more than 83.7% reported that their workload increased. Even though this might be caused by the learning curve related to the abrupt change in activities, the private Mexican university established different strategies to support its faculty in this and other wellbeing issues, and the analysis of their impact is the topic of another publication from this research group.

From the perspective of professors used to face-to-face modality, among the advantages involved in the online modality were saving time in transportation which also gave them opportunities to invest in their wellbeing like spending time with their families or having enough time to eat.

A relevant finding related to teaching from home was the direct correlation of the professors' performance with the appropriateness of a workspace and the consensus between family and work, in agreement with other authors (Ferri et al., 2020; Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe, 2020). Higher education institutions must consider this condition in case of future scenarios, which require similar shifts from a face-to-face modality to an online modality. Despite the increasing trends toward working from home, in this case, teaching from home did impede optimal performance of faculty.

Approximately 60% of the faculty participants preferred face-to-face over online instruction. Nevertheless, about 50% of them also recognized that digital environments involve multiple innovations. This abrupt change also represented a great opportunity for educational institutions and for all teachers, instructors, and professors to learn and use the best of both the face-to-face and online modalities in favor of a better education in emergencies.

6. Conclusion

Beyond the relevance of our findings, it is noteworthy to mention that for teachers and professors from other educational institutions conditions were different. In Mexico and Latin America, in general, a vast majority of them lacked digital competences. They also lacked competences for designing, implementing, and assessing online courses, with synchronous and asynchronous interactions (Area and Adell, 2021). Even more, the digital divide that is the differences in the access to electronic devices and a good internet connectivity, among learners and teachers worsened the situation and increased inequalities among the population (See et al., 2020; Area and Adell, 2021).

The evidence gives raise to the following highlights for higher education institutions and policy makers when considering education in emergency scenarios:

1. Most of the participants positively evaluated their performance in the online modality, which might be a consequence of the previous experience of the institution in innovation and in distance modality. Besides, it might be a consequence of the

freedom provided to the professors to decide on the optimal path for their courses.

2. Most participants preferred face-to-face over online modality. They recognized that the later diminished personal interactions and multiple distractors from learning. Since social relevance is an intrinsic motivation for learning, more strategies need to be explored to fully overcome this challenge.
3. With training in technological tools provided by the Institution, most of the participants promptly adapted their instruction to the online modality. Even though they also reported an increment in their workload, literature has reported a natural learning curve that might restore the equilibrium. The well-being of faculty must be the topic of further analysis.
4. Evidence demonstrated, through correlational statistics and classification techniques, that professors' performance depends on the conditions of their teaching space at home and the balance and harmony between their family issues and their job responsibilities. This provides valid arguments to promote the return of faculty to higher education facilities.
5. Evaluation and assessment methods in online modality require further research to ensure their effectiveness. Instead of adapting traditional evaluation methods into the digital environment, a new and improved assessment framework needs to be developed.
6. Evidence shows that the online modality represented a great opportunity to innovate. In order to enhance educational quality, institutions and educators should strive to integrate the most effective online and face-to-face teaching strategies, not solely in response to crises, but also in everyday academic contexts.

Our research on assessing professors' perceptions from a Mexican Higher education institution during the transitional phase to an Emergency Remote Teaching (ERT) amid COVID-19 pandemic holds relevance for other emergency situations that may arise in the future. As emergency situations are predicted to become more frequent in the coming years, our findings are of particular significance to higher education institutions and policymakers, who may draw upon the insights gained from our study to better prepare for and respond to such crises. Therefore, our research has broader implications for emergency planning and management to ensure education of future generations under any challenging situation.

Data availability statement

The datasets presented in this article are not readily available because the consent of the participants was obtained under the statement that the information collected during the study would be anonymized and kept confidential. Requests to access the datasets should be directed to CC-Z, claudia.camacho@tec.mx.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

MP-B, MC-C, and CC-Z contributed to conception and design of the study. CC-Z organized the database, performed the data mining, including text mining, and wrote the final draft of the manuscript. EA-R performed the statistical analysis. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to acknowledge the support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico in the production of this work.

References

- Alcántara Santuario, A. (2020). Educación superior y COVID-19: Una perspectiva comparada. In *Educación y pandemia: Una visión académica* (pp. 75–82). Universidad Nacional Autónoma de México, Instituto de Investigaciones sobre la Universidad y la Educación.
- Allen, J., Rowan, L., and Singh, P. (2020). Teaching and teacher education in the time of COVID-19. *Asia Pac. J. Teach. Educ.* 48, 233–236. doi: 10.1080/1359866X.2020.1752051
- Area, M., and Adell, J. (2021). Digital technologies and educational change. A critical approach. *Rev Iberoam. Sob. Calid. Efic. Cambio Educ.* 19, 83–96. doi: 10.15366/reice2021.19.4.005
- Bond, M., Bedenlier, S., Marín, V. I., and Händel, M. (2021). Emergency remote teaching in higher education: mapping the first global online semester. *Int. J. Educ. Technol. High. Educ.* 18:50. doi: 10.1186/s41239-021-00282-x
- Camacho-Zuñiga, C., Pego, L., Escamilla, J., and Hosseini, S. (2021). The impact of the COVID-19 pandemic on students' feelings at high school, undergraduate, and postgraduate levels. *Heliyon* 7:e06465. doi: 10.1016/J.HELIYON.2021.E06465
- Essa, A., and Mojarad, S. (2022). *Practical AI for Business Leaders, Product Managers, and Entrepreneurs*. 1st Edn. Berlin: De Gruyter.
- Ezra, O., Cohen, A., Bronshtein, A., Gabbay, H., and Baruth, O. (2021). Equity factors during the COVID-19 pandemic: difficulties in emergency remote teaching (ert) through online learning. *Educ. Inf. Technol.* 26, 7657–7681. doi: 10.1007/s10639-021-10632-x
- Ferri, F., Grifoni, P., and Guzzo, T. (2020). Online learning and emergency remote teaching: opportunities and challenges in emergency situations. *For. Soc.* 10:86. doi: 10.3390/soc10040086
- Flack, C. B., Walker, L., Bickerstaf, A., Earle, H., and Margetts, C. (2020). Educator perspectives on the impact of COVID-19 on teaching and learning in Australia and New Zealand. *Pivot Professional Learning*.
- Hayes, S. W., Priestley, J. L., Ishmakhametov, N., and Ray, H. E. (2020). "I'm not working from home, I'm living at work": perceived stress and work-related burnout before and during COVID-19. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/vnkwa
- Huang, R., Tlili, A., Chang, T.-W., Zhang, X., Nascimbeni, X., and Burgos, D. (2020). Disrupted classes, undisrupted learning during COVID-19 outbreak in China: application of open educational practices and resources. *Smart Learn. Environ.* 7. doi: 10.1186/s40561-020-00125-8
- Instituto Internacional de la UNESCO para la Educación Superior en América Latina y el Caribe (2020). COVID-19 and higher education: Today and tomorrow; impact analysis, policy responses and recommendations. UNESCO IESALC.
- Miguel Román, J. A. (2020). La educación superior en tiempos de pandemia: Una visión desde dentro del proceso formativo. *Rev. Latin. Estud. Educ.* 50, 13–40. doi: 10.48102/rlee.2020.50.ESPECIAL.95
- Moreno, J. M. (2021). Catastrophes of education or the endless debate; innovation as a fight against the elements. *Rev Iberoam. Sob. Calid. Efic. Cambio Educ.* 19, 97–110. doi: 10.15366/reice2021.19.4.006
- Penn, M. (2021). Statistics say large pandemics are more likely than we thought. DUKE Global Health Institute. Available at: <https://globalhealth.duke.edu/news/statistics-say-large-pandemics-are-more-likely-we-thought>
- Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037/10003-066X.55.1.68
- Sánchez Mendiola, M., Martínez Hernández, A. M. D. P., Torres Carrasco, R., de Agüero Servín, M., Hernández Romo, A. K., Benavides Lara, M. A., et al. (2020). Retos educativos durante la pandemia de COVID-19: Una encuesta a profesores de la UNAM. *Rev. Digit. Univ.* 21. doi: 10.22201/codeic.16076079e.2020.v21n3.a12
- Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., et al. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: a meta-analytic study. *Educ. Res. Rev.* 19, 158–172. doi: 10.1016/j.edurev.2016.07.003
- See, B. H., Wardle, L., and Collie, P. (2020). Teachers' wellbeing and workload during Covid-19 lockdown. Durham University Evidence Centre for Education.
- Sinclair, S., and Rockwell, G. (2016). Voyant tools. Available at: <http://voyant-tools.org/>
- Stewart, W. H. (2021). A global crash-course in teaching and learning online: a thematic review of empirical emergency remote teaching (ERT) studies in higher education during year 1 of COVID-19. *Open Praxis* 13:89. doi: 10.5944/openpraxis.13.1.1177
- Torani, S., Majd, P. M., Maroufi, S. S., Dowlati, M., and Sheikhi, R. A. (2019). The importance of education on disasters and emergencies: a review article. *IEHP*.
- Torras Virgili, M. E. (2021). Emergency Remote Teaching: Las TIC aplicadas a la educación durante el confinamiento por COVID-19. *Int J Technol Educ Innov* 7, 122–136. doi: 10.24310/innoeduca.2021.v7i1.9079
- Trust, T., and Whalen, J. (2020). Should teachers be trained in emergency remote teaching? Lessons learned from the COVID-19 pandemic. *J. Technol. Teach. Educ.* 28, 189–199.
- UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean, Economic Commission for Latin America and the Caribbean (2020). Education in the time of COVID-19. UNESDOC Digital Library. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000374075_eng
- United Nations (2022). The sustainable development goals report 2022.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.



OPEN ACCESS

EDITED BY

Kamal Kant Hiran,
Symbiosis University of Applied Sciences, India

REVIEWED BY

Sandra Sanchez-Gordon,
Escuela Politécnica Nacional, Ecuador
Abdallah Qusef,
Princess Sumaya University for Technology,
Jordan

*CORRESPONDENCE

Eva-Maria Schön
✉ eva-maria.schoen@hs-emden-leer.de

RECEIVED 19 April 2023

ACCEPTED 12 July 2023

PUBLISHED 27 July 2023

CITATION

Schön E-M, Neumann M, Hofmann-Stölting C,
Baeza-Yates R and Rauschenberger M (2023)
How are AI assistants changing higher
education? *Front. Comput. Sci.* 5:1208550.
doi: 10.3389/fcomp.2023.1208550

COPYRIGHT

© 2023 Schön, Neumann, Hofmann-Stölting,
Baeza-Yates and Rauschenberger. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)
(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.

How are AI assistants changing higher education?

Eva-Maria Schön^{1*}, Michael Neumann²,
Christina Hofmann-Stölting³, Ricardo Baeza-Yates⁴ and
Maria Rauschenberger⁵

¹Faculty of Business Studies, University of Applied Sciences Emden/Leer, Emden, Germany, ²Faculty of Business and Computer Science, Department of Business Information Systems, University of Applied Sciences and Arts Hannover, Hannover, Germany, ³Faculty of Business and Social Sciences, Department of Business, University of Applied Sciences (HAW) Hamburg, Hamburg, Germany, ⁴Institute for Experiential AI, Northeastern University, Boston, MA, United States, ⁵Faculty of Technology, University of Applied Sciences Emden/Leer, Emden, Germany

Context: Higher education is changing at an accelerating pace due to the widespread use of digital teaching and emerging technologies. In particular, AI assistants such as ChatGPT pose significant challenges for higher education institutions because they bring change to several areas, such as learning assessments or learning experiences.

Objective: Our objective is to discuss the impact of AI assistants in the context of higher education, outline possible changes to the context, and present recommendations for adapting to change.

Method: We review related work and develop a conceptual structure that visualizes the role of AI assistants in higher education.

Results: The conceptual structure distinguishes between humans, learning, organization, and disruptor, which guides our discussion regarding the implications of AI assistant usage in higher education. The discussion is based on evidence from related literature.

Conclusion: AI assistants will change the context of higher education in a disruptive manner, and the tipping point for this transformation has already been reached. It is in our hands to shape this transformation.

KEYWORDS

AI, ChatGPT, higher education, disruption, agile values

1. Introduction

The context of higher education is changing at an accelerating pace. During the COVID-19 pandemic, digital teaching became the new method of teaching. In addition, new learning concepts have evolved, and collaborative technologies have spread. Innovative teaching concepts have been explored, such as gamification frameworks in learning environments (Rauschenberger et al., 2019), agile approaches (Neumann and Baumann, 2021; Schön et al., 2022), or the use of emerging technologies (e.g., a robot that serves as a teaching assistant and scrum master; Buchem and Baecker, 2022). Nowadays, emerging technologies, which include AI tools such as ChatGPT, are changing the context of higher education in a disruptive manner (Haque et al., 2022). ChatGPT is a large language model based on GPT-3 and was released by the company OpenAI in November 2022 (OpenAI, 2022). The AI chatbot provides real-time communication for users prompting their requests. The quality of ChatGPT's *natural speaking* answers marks a major change in how we will use AI-generated information in our day-to-day lives and has the potential to completely alter our interaction with technology (Aljanabi et al., 2023). More recently, Google launched *Bard* (which is currently only available for a small group of external testers), an AI chatbot

competitor to ChatGPT that is tailored to search tasks and can even use current information on the web to answer questions (The decoder, 2021).

There have been many discussions regarding the potential impact of ChatGPT, with some viewing it as a disruptive technology (Haque et al., 2022; Rudolph et al., 2023). Some even believe it should be prohibited due to the change it may bring. However, ChatGPT is still in its infancy and makes mistakes (Gao et al., 2022). But, at the same time, it has great potential for the future. For instance, it could be used for voice user interfaces to overcome issues with response behavior or response quality (Klein et al., 2021). AI assistants can support the creation of new ideas or help in terms of automating tasks. Hence, AI assistants will change the way people work.

This development also has implications for the context of higher education. On the one hand, it poses challenges such as unknown handling by students, more time-consuming assessments, and unknown potential. On the other hand, it offers opportunities such as increased individual intelligent tutoring systems (ITS) and creativity engagement (Neumann et al., 2023). One of the challenges with the output of ChatGPT in the context of higher education is that established control structures (such as plagiarism checkers or AI detection tools) are not able to detect whether the text is generated by an AI or a human (Gao et al., 2022). As a result, we will have a change in the way we conduct and evaluate exams. In addition, we must consider creating good practices for using ChatGPT in a responsible and ethical manner, e.g., Atlas (2023). These examples will not be the only changes. Equally important is a change at the cultural level toward a student-centered approach and value-based learning (Schön et al., 2022). However, it is obvious that the context of higher education is changing. Thus, we need discussion and guidance on how to deal with such emerging technologies in order to actively shape this transformation.

This paper examines the following research questions (RQ):

RQ1: What is the impact of AI assistants in the context of higher education?

RQ2: How can higher education institutions adapt to the changes brought by AI assistants?

To answer RQ1, we developed a conceptual structure that highlights how AI assistants will change the context of higher education. The aim of our conceptual structure is to formalize ongoing activities in terms of the changes brought by generative AI assistants such as ChatGPT. The conceptual structure distinguishes between humans, learning, organization, and disruptor. It allows us to have a guided discussion regarding the implications of AI assistant usage in higher education. We enriched the discussion with evidence from related literature and our own experience with different use cases using the chatbot ChatGPT (Schön et al., 2023). To answer RQ2, we present lessons learned from the agile community that allow us to outline good practices for adapting to change.

This paper is structured as follows: Section 2 briefly summarizes the related work. Section 3 presents our conceptual structure for the usage of AI assistants in the context of higher education. Section 4

outlines the implications of AI assistant usage in higher education and discusses the changes in terms of humans, learning, and organizational aspects. Section 5 explores ethical issues regarding the use of AI assistants in the context of higher education and presents lessons learned from the agile community concerning transformation toward value-based working. This paper closes with a conclusion and future work in Section 6.

2. Related work

AI assistants have become more relevant in recent years and have reached the context of higher education, which is shown by the increasing number of publications and literature reviews conducted on this topic (Ouyang et al., 2022). Thus, we looked for secondary studies related to AI in higher education using Google Scholar. Secondary studies examine all primary studies related to a particular research question or topic, with the goal of synthesizing the evidence related to that question. The latest reviews concerning AI in higher education are briefly summarized below.

Alam and Mohanty (2022) surveyed existing literature in a systematic manner with the objective of identifying and examining the ethical considerations, challenges, and potential threats associated with using AI in higher education as well as exploring the potential uses of AI. They grouped their results into four categories: *intelligent tutoring systems*, *personalization and adaptive systems*, *evaluation and assessment*, and *prediction and profiling*. The authors state that their research reveals a lack of critical thinking regarding the challenges and potential threats of using AI in higher education.

Another systematic review by Ouyang et al. (2022) focuses on AI in online higher education. Their research aims to examine the various purposes for which AI is applied, the AI algorithms utilized, and the outcomes produced by AI techniques in online higher education. In terms of teaching, the authors found that AI applications are used to predict learning status, performance or satisfaction, resource recommendation, automatic assessment, and improvement of the learning experience. The authors claim that AI has been a crucial aspect of education from the perspectives of instructors, learners, and administrators, with the ability to create both opportunities and challenges in the transformation of higher education.

Since the release of ChatGPT in November 2022, there have already been several publications on the AI assistant. Rudolph et al. (2023) surveyed the existing literature regarding ChatGPT and higher education and found some peer-reviewed articles and preprints, which they included in their review. The authors also performed queries with ChatGPT. Their article presents the strengths and limitations of ChatGPT and discusses the implications of ChatGPT for higher education concerning *student-facing AI applications*, *teacher-facing AI applications*, and *system-facing AI applications*. Moreover, they offer recommendations for handling ChatGPT in higher education. The authors categorize ChatGPT as an AI-powered writing assistant. They conclude that ChatGPT can be beneficial for providing conceptual explanations and applications but cannot create content that requires higher-order thinking (such as critical or analytical thinking).

The related work provides evidence that allows us to create our conceptual structure and discuss the impact of AI assistants in higher education. Compared to our research, the paper of Rudolph et al. (2023) presents a good overview of current developments regarding ChatGPT. However, our conceptual structure goes one step beyond and allows us to structure the knowledge and discussion on a meta-level. Summarizing our findings, we can say that the quality of AI tools is rapidly improving. As outlined in the introduction, the quality of the text generated by ChatGPT is impressive and will change the context of higher education in a disruptive manner. Thus, we have reached a tipping point at which change has been initiated on several levels, marking a major transformation. The question now is about how we want to deal with this transformation. Hence, the following section presents a conceptual structure that outlines the changes that AI assistants will bring to the context of higher education, followed by a discussion of the implications.

3. Developing a conceptual structure for AI assistants in higher education

A conceptual structure is a way to describe the organization and connections between various components of a specific system, similar to a meta-model, which goes beyond and creates the basis for a language used for creating models (Escalona and Koch, 2007; Schön et al., 2019). Moreover, a conceptual structure provides a type of classification and allows us to make inferences and predictions based on selected information (Medin, 1989). It facilitates a shared understanding of a particular problem area and provides an abstract perspective on the problem. We want to answer our RQ1: *What is the impact of AI assistants in the context of higher education?* with this conceptual structure (see Figure 1) and the following discussion of implications in the Section 4.

3.1. Developing the conceptual structure

We used a formalized approach to develop our conceptual structure. In the beginning, we started with a review of gray literature related to ChatGPT (Neumann et al., 2023). Moreover, we tested different use cases using the chatbot ChatGPT (Schön et al., 2023). Next, we had several discussions with lecturers and researchers to better understand how AI assistants will change higher education. We used a Miro board to visualize and capture our discussions and thoughts. We then started to create the first version of our conceptual structure based on our preliminary work and related literature. We used a UML (Object Management Group, 2017) notation since it provides a formal representation method that is commonly used. The conceptual structure was refined over eight iterations through discussions among the authors. It is presented in Figure 1.

3.2. Conceptual structure for AI assistants in higher education

This section presents our conceptual structure for AI assistants in higher education and discusses the concepts and relations between the classes. We present in our conceptual structure (see Figure 1) four different areas (see the color encoding) related to the change in higher education due to AI assistants:

Humans (yellow): classes *Lecturer*, *Student*

Learning (green): classes *Learning Experience*, *Learning Assessment*, *Module*

Organization (blue): classes *University*, *Regulation*

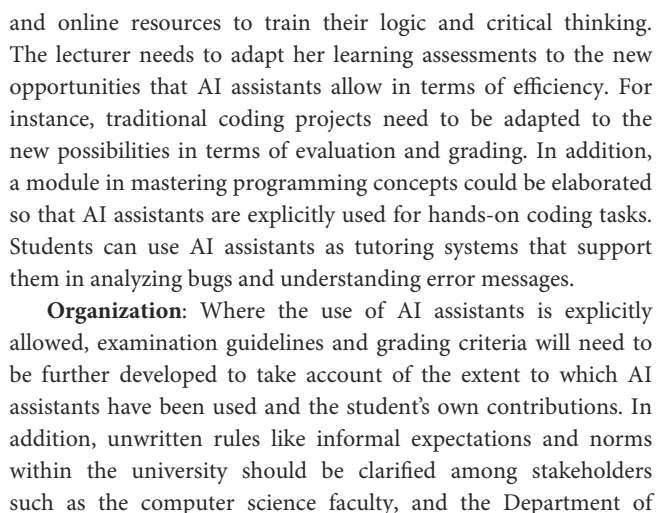
Disruptor (red): class *AI Assistant*

The conceptual structure contains the class **Lecturer** with the following attributes: *topic*, *skill*, *competency*, *mindset*, and *way of working*. **Lecturers** interact with **Students** who are described by a *degree program*, *skills*, *competency*, *mindset*, and *learning style*. Both classes are related to the concept of **Humans**. Moreover, the concept of **Learning** comprises the classes **Learning Experience**, **Learning Assessment**, and **Module**. **Students** have **Learning Experience** which differs in terms of *prior knowledge*, *mindset/values*, *condition*, *skills*, and *tools*. **Lecturers** are responsible for **Learning Assessments** that vary in terms of *type*, *evaluation*, and *grading*. **Learning Assessments** are part of **Modules** covering *didactic goals*, *teaching concepts*, *didactic methods*, *learning assessments*, and *course material*. **Modules** are taught by **Lectures** and influenced by **Regulations**. **Regulations** are determined by *examination regulations*, *degree program regulations*, or *unwritten rules*. Together with the class **University**, they represent the concept of **Organization**. **University** has *faculty*, *department*, *institution*, *degree program*, and *university committee*. In the middle of Figure 1, the concept of **Disruptor** is placed as the class **AI Assistant**. This one is highlighted in red because it changes the whole context of higher education. As outlined above, when **Lecturers** and **Students** use **Assistants**, this will change the **Learning Experience**, **Learning Assessment**, **Module**, and **Regulation**. We will discuss these changes in the next section and outline the impact of using **AI Assistants** in higher education.

In the following, we present an example of how our conceptual structure can lead a discussion regarding the implications of AI assistant usage in higher education in reality:

Human: The lecturer in this example is mainly responsible for the topic of artificial intelligence. Her skill is programming and her competencies are focused on research in machine learning. Her mindset is innovative and adaptable to new circumstances and she engages in collaborative teaching methods. Usually, the lecturer interacts with students in computer science, who have special skills in problem-solving and coding. The students' competencies are in understanding algorithms. Their mindset is curious and they are eager to learn. Most of them are visual learners.

Learning: The student's learning experience differs in terms of their knowledge with the understanding of basic programming concepts. They learn in a mix of classroom-based or online learning conditions. In doing so, they use programming software



This example shows that by applying our conceptual structure, a structured representation of the changes to be considered those trigger AI assistants is possible. Finally, it can also be applied to other domains and study programs.

4. Implications of AI assistants in higher education

This section will discuss the impact of AI assistants according to the four previously-defined concepts (see Section 3): *Humans*, *Learning*, *Organization*, and *Disruptor*. The conceptual structure (see Figure 1) uses the relationships highlighted in red to show the extent to which AI assistants may change the context of

higher education. We focus on how areas of higher education are being transformed by AI assistants using the example of ChatGPT. Therefore, we use related work to enrich the following discussion of implications and execute tests with ChatGPT.¹

4.1. Humans

Students can use AI assistants to identify strengths or gaps in their knowledge and to receive personalized feedback on their learning progress, or work results (Zawacki-Richter et al., 2019). Thus, they are individually supported in the development of their competencies. ChatGPT may help one to improve their academic research and writing skills. It can summarize papers, extract key facts, and even provide citations and references. The tool can also assist (not replace) academic writing skills by generating essays or parts of essays for papers, dissertations, or similar work. Furthermore, ChatGPT can give feedback and correct text passages (Aljanabi et al., 2023). There are various research papers in which ChatGPT is used to write literature review articles with promising results (e.g., Aydin, 2022). This implies that AI assistants will play an essential role in the field of research and writing to support academics.

Lecturers may use AI assistants to reduce their workload by automating assessment, administration, and feedback mechanisms (Rudolph et al., 2023). In particular, the time saved by automating assessment allows Lecturers to focus on empathetic human teaching (Zawacki-Richter et al., 2019). In addition, lecturers can use AI assistants for lesson planning by having them create a course syllabus with short descriptions of the topics (Kasneci et al., 2023). Moreover, AI assistants can help lecturers to create materials for different learning levels. They may also transfer solution examples (e.g., from one programming language to another) to save time. Test queries show that different examples can be generated and easily transformed by ChatGPT to suit all levels, from beginners to experts (Schön et al., 2023).

Both students and lecturers need to develop competencies so that they can use AI assistants effectively. At the same time, it is important to establish a proper mindset and raise awareness about the ethical aspects of AI assistant usage. Because ChatGPT generates near-perfect natural speech answers, a human may think “*That must be correct.*” ChatGPT also has other limitations: generated answers can be too short, misinterpreted, not understandable for students, or wrong (Gao et al., 2022; Qadir, 2022; Rudolph et al., 2023; Schön et al., 2023). Thus, humans must always evaluate the quality of ChatGPT’s answers. Misinterpretation of queries can lead the AI assistant to learn the individual wording of students and lectures (like predictive text on a smartphone) to avoid misinterpretation in the future. A solution could be for lecturers to provide videos or audio recordings of their own lectures and then query exam questions from the AI assistants. We should keep in mind that AI assistants should be controlled by humans and that even the best AI can make mistakes (Zawacki-Richter et al., 2019; Perry et al., 2023).

4.2. Learning

AI assistants are revolutionizing higher education, creating both opportunities and challenges for enhancing the quality of higher education and improving the learning experience (Ouyang et al., 2022).

In particular, since AI assistants can serve as intelligent tutoring systems (ITS), they are changing how students learn (Zawacki-Richter et al., 2019; Fauzi et al., 2023; Rudolph et al., 2023). For instance, ChatGPT can serve as an AI-powered writing assistant and will therefore bring innovation to certain types of tasks. Lecturers can provide students with learning material that the students can then work through at their own pace since AI assistants can give them feedback. This would change the learning experience. Students can also use AI assistants to create individualized learning paths and personalized learning instructions according to their prior knowledge, conditions, and pace. Hence, students can interact with AI assistants such as ChatGPT and engage with content that is new to them and fits their needs. Especially with regard to large-scale lectures and massive open online courses, AI assistants have the potential to create individual learning experiences (Winkler and Soellner, 2018).

For exam preparation, AI assistants can also be used as intelligent tutoring systems (ITS). For example, questions about texts and other learning materials can be generated as a *mock exam*. ChatGPT is able to generate multiple-choice questions, quizzes, open questions, and much more. It is also conceivable that the AI assistant can not only provide the solution to a complex problem (e.g., a math problem) but also an individualized explanation of the solution process.

When implementing software code, computer science students can use AI-generated tests (e.g., unit tests) to test their own program codes. However, there are some limitations since AI-generated code is significantly less secure (Perry et al., 2023). Still, when asked for code, ChatGPT does sometimes unintentionally suggest secure improvements (Schön et al., 2023). In addition, ChatGPT could potentially supplement or replace Google search or communities such as *Stackoverflow*.² This is because ChatGPT responds in a matter of seconds, whereas communities with actual humans need hours, days, or even weeks to answer.

AI assistants are already changing learning assessments. Assessment types that are generic and could be created by a human or an AI assistant should be avoided. Instead, assessments should be designed to develop students’ creative and critical thinking skills (Rudolph et al., 2023).

These assessment types could cover presentations as well as multimedia content (such as videos, websites, or animations). Another type could be a *stealth assessment*: a continuous, integrated, and inconspicuous method of evaluation that takes place in various forms (such as serious games, simulations, virtual labs, or forums). It involves collecting data on student performance while they engage in tasks (Caspari-Sadeghi, 2023).

When a large number of students need to be assessed, automated assessment plays an important role. Automating

¹ Examples of our test prompts can be found in Schön et al. (2023).

² Available online at: <https://stackoverflow.com/> (accessed February 10, 2023).

assessment also allows the students to conduct the assessment whenever they are ready to do so. Students may benefit from flexible timed exams in a variety of subjects such as math, as it allows them to take the exam when they feel ready.

Tasks that are comparable and coordinated according to content, difficulty, and level of competence can be generated automatically. Lecturers can also use AI assistants to grade exams, including giving individual feedback regarding strengths and weaknesses for various types of assessments (such as essays, research papers, and written exams, Kasneci et al., 2023).

The use of AI assistants affects instructional design and implementation through various educational perspectives, thus having a significant impact (Ouyang et al., 2022). The increasing digitalization requires skills such as working in an interdisciplinary team and self-organization. Modules should cover those aspects in terms of didactic goals and didactic methods. An example of how those competencies could be achieved is the experience of working in an agile team. Therefore, no-code platforms in combination with agile methods and agile practices are used (Lebens and Finnegan, 2021). AI assistants can support students who are not primarily studying computer science to gain this experience. With the increasing number of low-code platforms (e.g., Salesforce³) and ChatGPT, more and more people are able to create source code. This will bring more change to all degree programs since other competencies will be in demand. For example, business students may soon automate their work themselves instead of having a programmer for each task. This is a controversial societal issue that needs to be addressed.

Using an AI assistant requires certain competencies, just as using the web requires media competencies. Therefore, students need to learn how to use AI assistants and develop competence for different tasks according to their study program (e.g., informatics vs. business vs. social pedagogy). The current question is how and what we want to teach students, as we do not yet have much experience with AI assistants in daily use. Thus, research needs to be done. In addition, students are going into industries; therefore, we need to determine what companies need from their future employees regarding AI assistant usage competencies.

As for the concept of humans, the same is valid for concept learning: we need to be aware of the kinds of mistakes ChatGPT or any other AI assistant makes. The competence for ethical use needs to be conveyed. In addition, students must develop a mindset of wanting to learn due to intrinsic motivation of mastery and purpose.

4.3. Organization

As outlined above, AI assistants will have an impact on regulations. For instance, they will change the learning assessments in terms of types and evaluations, thus impacting examination regulations and degree program regulations. For instance, automatic assessment is one of the ways AI is already being used in higher education (Zawacki-Richter et al., 2019; Ouyang et al., 2022). Lecturers are concerned that ChatGPT will

change the process of writing as we know it and that traditional assessment types (such as essays and take-home exams) must be reformed (Rudolph et al., 2023). ChatGPT is capable of creating essays in just a few seconds. Established control mechanisms, such as plagiarism checkers or AI output detectors, cannot reliably and consistently identify generated texts (Gao et al., 2022).

If lecturers use AI assistants to generate exams and students use AI assistants to answer exam questions, then AI assistant usage will have reached a point of absurdity. Since this would mean the AI generates and completes the exams, it would no longer represent the learning level of students. Hence, we need to discuss limits regarding AI assistant usage or other approaches to assess the learning levels of students. Given the possibilities, one must consider whether certain forms of learning assessments, such as term papers and online exams, still make sense in the modern day (Susnjak, 2022). Additionally, new forms of learning assessments in which ChatGPT is explicitly used or does not provide any added value should be developed. For example, examinations could be designed to assess higher levels of competence, such as critical thinking or problem-solving skills (Cotton et al., 2023).

Despite the implications of AI assistants for regulations, these tools can deliver value-added services to other areas of universities. AI assistants can be used for profiling and prediction. Such AI assistants rely on learner profiles or models to make predictions, such as the risk of a student dropping out of a course or the likelihood of their admission to a program. This information can then be used to provide timely support, feedback, and guidance in content-related matters during the learning process (Zawacki-Richter et al., 2019). Another example of value-added services is that AI assistants can support course guidance services, such as answering questions about specific courses at a university. However, it can also provide individual advice for courses of study based on individual skills and prior experiences. AI assistants can also support the office of student affairs or other offices concerned with study organization, matriculation, certificates, and FAQs, although many legal and ethical questions remain.

When proven control mechanisms no longer work, and types of learning assessments change, new ways of dealing with deception and evaluation must be found. In this context, it is important to develop rules for dealing with AI assistants that all parties involved can understand and follow. There are still many open questions, such as *Is it plagiarism if AI writes an essay?* or *Who is the author of AI-generated texts?* Because of this, a university should coordinate legal opinions and provide clear information to students and staff (Ruhr University Bochum, 2023). The legal opinion is intended to provide guidance to universities in North Rhine-Westphalia (Germany) on how issues related to ChatGPT and similar programs should be handled regarding copyright and examination laws. However, some copyright issues can only be resolved with the providers of the AI assistants since they train the models with data, and it is not always transparent whether all training data was used legally (Kasneci et al., 2023).

5. Discussion

So far, we have outlined the impact of AI assistants in the context of higher education. However, there is another important topic that needs to be addressed regarding the use of AI assistants.

³ Available online at: <https://www.salesforce.com/eu/> (accessed February 10, 2023).

As with other emerging technologies, we need to be aware of ethical issues especially when technology is used to support decision-making. Therefore, we present our concerns below. In addition, this section presents lessons learned from the agile community that are relevant to transformation in the context of higher education and can support leaders.

5.1. Ethical issues

There are several ethical issues related to ChatGPT, which is based on large language models (LLMs). The main ones, in order of importance, are:

- **Discrimination:** LLMs learn the human biases that appear in the training texts, so they can sometimes generate biased, unfair, or discriminatory answers that may psychologically harm some readers. In addition, since answers depend on the prompt language, there is also a quality bias according to the language used (that is, for the same prompt, the answer may be different for different languages).
- **No attribution:** Current LLMs cannot give attribution to the texts that are used to generate an answer. Hence, ownership, copyrights, and other intellectual rights are not protected.
- **Weak and arrogant character:** ChatGPT trusts humans, and, in that sense, it is very naive and can be easily manipulated (e.g., tell me why drinking chlorine is good). On the other hand, it always answers with confidence, even when it is wrong. So, people may believe it unless they can easily check the facts.
- **Consent and privacy:** LLMs are mainly trained using web documents that may have usage restrictions that were not respected. Even if there are no restrictions, there is never explicit consent to allow these documents to be used as training data. As a result, the generated answers may reveal private information, violating personal or institutional privacy.

The ownership and copyright of generated text of ChatGPT are unclear, and one could believe a text belongs to the person who prompted it. Otherwise, one could argue that the AI-generated output belongs to the AI and needs to be cited. Either way, we as a society or institutions need to decide how to handle AI-generated output.

5.2. Lessons learned from the agile community

As described above, several AI assistants exist and are currently used by students and lecturers for various use cases. However, with the release of ChatGPT and the immense interest in that topic, a disruptive change has already begun. Thus, one may predict that tools such as ChatGPT will not disappear. Furthermore, we expect the release of further AI tools [for a list of AI-based tool examples, see (Schön et al., 2023)] and their integration into the existing tool landscape (e.g., the Microsoft Office Suite). This situation

may be described as a tipping point from a change management perspective, as disruptive change has already begun.

Disruptive changes are not a new phenomenon that only occurs due to the release of AI tools. Moreover, we know such situations from various examples: the rising dynamic of the markets, including the challenges to meet customer needs at an accelerating pace, navigating in uncertain and fast-paced environments, or the fundamental changes in work organization due to the COVID-19 pandemic, just to name a few. The agile community is used to disruptive changes, and we know from several aspects (e.g., agile transformation) that such rapid transitions come with various challenges (e.g., Dikert et al., 2016; Schön et al., 2017; Karvonen et al., 2018; Strode et al., 2022). So, what can we learn from the agile community to support this transformation in the context of higher education? The objective of the following discussion is to answer our RQ2: *How can higher education institutions adapt to the changes brought by AI assistants?*

First, we want to point to one major challenge, which is well-known in the area of agile transformation and has gained more and more research interest in recent years: the need for a cultural change (Sidky et al., 2007; Kuchel et al., 2023). We know that the interplay between technical and cultural agility (also known as *being* vs. *doing* agile) is of high importance when using or introducing agile methods and practices (Diebold et al., 2015; Küpper et al., 2017). Hence, the fit of cultural values in an organization and the underlying values and principles of agile (Schwaber and Sutherland, 2020; Beck et al., 2021) requires a cultural change.

In-depth knowledge from the area of agile transformation, particularly regarding cultural aspects, may support us in tackling the upcoming challenges and promoting the potential of AI tools for higher education. From our point of view, it is important to establish a culture of trust, especially between lecturers and students. Established control mechanisms such as plagiarism scanners are now useless. Currently, there are no reliable technical methods for determining whether a text or other content was generated by an AI assistant such as ChatGPT or by a human being. Thus, we argue that there is a need for a cultural shift toward a value-based learning approach that focuses on a trustful learning environment. This requires new competencies of students and lecturers (e.g., self-organization or adaption of the learning process according to continuous feedback, Schön et al., 2022) as well as new learning assessments. We also see the need for a defined set of values and principles surrounding this topic. This set would provide a foundation for a value-based learning approach regarding the upcoming aspects of AI assistant usage.

Summarizing this discussion, we point to the aspect of *being* vs. *doing* agile WRT with regard to the integration of AI assistants into the higher education context. As AI assistants are available and already used by both students and lecturers, the technical facet (*doing*) is covered. However, as described above, the facet of values and principles (*being*) needs our focus, e.g., through discussion of how we can enable a new mindset in the higher education context. By finding comprehensive solutions that address both facets, we can create a future-oriented, sustainable, and resilient educational environment, which may be described as an agile higher educational context.

5.3. Limitations

Although our study was designed and conducted according to established guidelines, it is important to consider certain limitations.

5.3.1. Construct validity

We developed our conceptual structure for AI in higher education inductively in several iterations based on existing literature and discussions among researchers. Since this research is about an emergent technology, it cannot cover deeply unresolved issues such as the ethical implications of the use of AI assistants in the context of higher education. However, with our conceptual structure, we want to outline the open issues, that need to be addressed in future research.

5.3.2. Internal validity

We outlined the connection between the concepts and relationships of our conceptual structure and the existing literature in Section 4. One potential concern is the possibility of confounding variables, which could have influenced the observed relationships between the concepts. We tried to avoid this through intensive, structured discussions among the group of authors.

5.3.3. External validity

The aim of our conceptual structure is to formalize ongoing activities in terms of the changes brought about by generative AI assistants such as ChatGPT. As an emerging technology, we cannot be sure that we have considered all aspects. Therefore, we need to conduct case studies in the future to observe and better understand the changes. Our findings are currently applicable to similar AI assistants such as chatbots from other providers, and also to other areas such as image generation or speech generation (see [Figure 1](#), type of AI Assistant).

6. Conclusion and future work

This paper presents a conceptual structure that highlights the changes that AI assistants bring to the context of higher education. Our conceptual structure was developed by means of a literature review and extensive discussion among researchers and lecturers. The conceptual structure covers the following concepts: Humans, Learning, and Organization. Furthermore, we discuss the implications that AI assistants have for the context of higher education. The implications comprise changes addressing: (1) *how lecturers and students teach and learn*, (2) *the competencies that both need for the ethical and technical usage of AI assistants*, (3) *the learning experience*, (4) *the evolution of learning assessments and grading*, (5) *changes to regulations*, and (6) *topics that must be addressed by the organization*. Moreover, we presented critical aspects such as ethical issues as well as changes to the value system and mindset that must be guided by the people who are involved in the transformation process (e.g., students, lecturers). Therefore, we presented lessons learned from the agile community.

Technological innovations, such as the digital calculator or search engines, have challenged common practices in research and teaching, as AI assistants are doing today. However, the urgency with which we must act is greater for AI assistants, as the tool has already been used by many people in a short amount of time. We need to examine our processes and adapt them to new circumstances, addressing the concepts Humans, Learning, and Organization. At the same time, we need to consider a shift in value systems toward a value-based learning approach that requires new competencies of lecturers and students. Only then can we definitely take advantage of AI assistants in higher education.

In future work, we want to investigate the impact of AI assistants in higher education through empirical study. For instance, we started case studies that investigate the changes that are brought by AI assistants to higher education. Since this research is about an emergent technology, we have to observe and analyze this phenomenon in the long term to cover deeply unresolved issues such as the ethical implications of the use of AI assistants in higher education. In addition, we want to shape the transformation process and clarify what should be discussed regarding the ethical use of AI assistants and the change toward a value-based learning approach.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://figshare.com/s/5a3297b38934470b6c69>.

Author contributions

E-MS, MN, CH-S, RB-Y, and MR: conceptualization, methodology, writing—original draft, review, and editing. All authors contributed to the article and approved the submitted version.

Funding

The article processing charges for this article are supported by the University of Applied Sciences Emden/Leer.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Alam, A., and Mohanty, A. (2022). "Foundation for the future of higher education or 'misplaced optimism'? Being human in the age of artificial intelligence," in *Communications in Computer and Information Science*, eds M. Panda, S. Dehuri, M. R. Patra, P. K. Behera, G. A. Tshirintzis, S.-B. Cho, and C. A. Coello (Springer Science and Business Media Deutschland GmbH), 17–29. doi: 10.1007/978-3-031-23233-6_2
- Aljanabi, M., Ghazi, M., Ali, A. H., and Abed, A. (2023). ChatGPT: open possibilities. *Iraqi J. Comput. Sci. Math.* 4, 62–64. doi: 10.52866/20ijscm.2023.01.01.0018
- Atlas, S. (2023). *ChatGPT for Higher Education and Professional Development: A Guide to Conversational AI*. DigitalCommon@URI.
- Aydin, Ö. (2022). *OpenAI ChatGPT Generated Literature Review: Digital Twin in Healthcare*. Available online at: https://papers.ssrn.com/sol3/papers.cfmabstract_id=4308687 (accessed February 10, 2023).
- Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., et al. (2021). *Agile Manifesto*. Available online at: <https://agilemanifesto.org/> (accessed February 10, 2023).
- Buchem, I., and Baecker, N. (2022). "Nao robot as scrum master: Results from a scenario-based study on building rapport with a humanoid robot in hybrid higher education settings," in *Training, Education, and Learning Sciences*, Vol. 59 (AHFE International), 65–73.
- Caspari-Sadeghi, S. (2023). Learning assessment in the age of big data: learning analytics in higher education. *Cogent Educ.* 10, 2162697. doi: 10.1080/2331186X.2022.2162697
- Cotton, D. R., Cotton, P. A., and Shipway, J. R. (2023). *Chatting and Cheating: Ensuring academic integrity in the era of ChatGPT*. Available online at: <https://edarxiv.org/mrz8h/download?format=pdf> (accessed February 10, 2023).
- Diebold, P., Küpper, S., and Zehler, T. (2015). "Nachhaltige agile transition: Symbiose von technischer und kultureller agilität (sustainable agile transition: Symbiosis of technical and cultural agility)," in *Projektmanagement und Vorgehensmodelle 2015*, eds M. Engstler, M. Fazal-Baqaie, E. Hanser, M. Mikusz, and A. Volland (Bonn: Gesellschaft für Informatik e.V.), 121–126.
- Dikert, K., Paasivaara, M., and Lassenius, C. (2016). Challenges and success factors for large-scale agile transformations: a systematic literature review. *J. Syst. Softw.* 119, 87–108. doi: 10.1016/j.jss.2016.06.013
- Escalona, M. J., and Koch, N. (2007). "Metamodeling the requirements of web systems," in *Web Information Systems and Technologies*, eds J. Filipe, J. Cordeiro, and V. Pedrosa (Berlin/Heidelberg: Springer), 267–280.
- Fauzi, F., Tuhuteru, L., Sampe, F., Ausat, A. M. A., and Hatta, H. R. (2023). Analysing the role of chatGPT in improving student productivity in higher education. *J. Educ.* 5, 14886–14891. doi: 10.31004/joe.v5i4.2563
- Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., et al. (2022). Comparing scientific abstracts generated by chatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *bioRxiv*. doi: 10.1101/2022.12.23.521610
- Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., and Ahmad, H. (2022). I think this is the most disruptive technology: Exploring sentiments of ChatGPT early adopters using twitter data. *arXiv[Preprint]*. arXiv: 2212.05856. doi: 10.48550/arXiv.2212.05856
- Karvonen, T., Sharp, H., and Barroca, L. (2018). "Enterprise agility: why is transformation so hard?" in *Agile Processes in Software Engineering and Extreme Programming*, eds J. Garbajosa, X. Wang, and A. Aguiar (Cham: Springer International Publishing), 131–145.
- Kasneji, E., Sessler, K., Kuechemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Diff.* 103, 102274. doi: 10.1016/j.lindif.2023.102274
- Klein, A. M., Rauschenberger, M., Thomaschewski, J., and Escalona, M. J. (2021). Comparing voice assistant risks and potential with technology-based users: a study from Germany and Spain. *J. Web Eng.* 7, 1991–2016. doi: 10.13052/jwe1540-9589.2071
- Kuchel, T., Neumann, M., Diebold, P., and Schön, E.-M. (2023). "Which challenges do exist with agile culture in practice?" in *SAC '23: Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing* (Tallinn), 1018–1025. doi: 10.1145/3555776.3578726
- Küpper, S., Kuhrmann, M., Wiatrok, M., Andelfinger, U., and Rausch, A. (2017). "Is there a blueprint for building an agile culture?" in *Projektmanagement und Vorgehensmodelle*, eds A. Volland, M. Engstler, M. Fazal-Baqaie, E. Hanser, O. Linssen, and M. Mikusz (Bonn: Gesellschaft für Informatik), 111–128.
- Lebens, M., and Finnegan, R. (2021). "Using a low code development environment to teach the agile methodology," in *Agile Processes in Software Engineering and Extreme Programming*, eds P. Gregory, C. Lassenius, X. Wang, and P. Kruchten (Cham: Springer International Publishing), 191–199.
- Medin, D. L. (1989). Concepts and conceptual structure. *Am. Psychol.* 44, 1469–1481.
- Neumann, M., and Baumann, L. (2021). "Agile methods in higher education: Adapting and using eduScrum with real world projects," in *Proceedings of the 2021 IEEE Frontiers in Education Conference (FIE)* (Lincoln, NE: IEEE), 1–8. doi: 10.1109/FIE49875.2021.9637344
- Neumann, M., Rauschenberger, M., and Schön, E.-M. (2023). "We Need to Talk About ChatGPT: The Future of AI and Higher Education. Hochschule Hannover.
- Object Management Group (2017). *Unified Modeling Language Version 2.5.1*. Technical report, Object Management Group.
- OpenAI (2022). *ChatGPT: Optimizing Language Models for Dialogue*. Available online at: <https://openai.com/blog/chatgpt/> (accessed February 10, 2023).
- Ouyang, F., Zheng, L., and Jiao, P. (2022). Artificial intelligence in online higher education: a systematic review of empirical research from 2011 to 2020. *Educ. Inform. Technol.* 27, 7893–7925. doi: 10.1007/s10639-022-10925-9
- Perry, N., Srivastava, M., Kumar, D., and Boneh, D. (2023). *Do Users Write More Insecure Code With AI Assistants?* Available online at: <https://anonymous.4open.science/r/ui> (accessed February 10, 2023).
- Qadir, J. (2022). *Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education*. doi: 10.36227/techrxiv.21789434.v1
- Rauschenberger, M., Willems, A., Ternieden, M., and Thomaschewski, J. (2019). Towards the use of gamification frameworks in learning environments. *J. Interact. Learn. Res.* 2019, 147–165.
- Rudolph, J., Tan, S., and Tan, S. (2023). ChatGPT: bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* 6, 343–363. doi: 10.37074/jalt.2023.6.1.9
- Ruhr University Bochum (2023). *Bochumer Projekt schafft Klarheit zu KI-Tools für NRW-Hochschulen (Bochum Project Provides Clarity on AI Tools for NRW Universities)*.
- Schön, E.-M., Buchem, I., and Sostak, S. (2022). "Agile in higher education: how can value-based learning be implemented in higher education?" in *Proceedings of the 18th International Conference on Web Information Systems and Technologies (WEBIST 2022)* (Valletta), 45–53. doi: 10.5220/0011537100003318
- Schön, E.-M., Neumann, M., Hofmann-Stölting, C., Baeza-Yates, R., and Rauschenberger, M. (2023). *Research protocol Chat GPT XP*. Available online at: <https://figshare.com/s/5a3297b38934470b6c69> (accessed July 17, 2023).
- Schön, E.-M., Sedeño, J., Mejías, M., Thomaschewski, J., and Escalona, M. J. (2019). A metamodel for agile requirements engineering. *J. Comput. Commun.* 7, 1–22. doi: 10.4236/jcc.2019.72001
- Schön, E.-M., Thomaschewski, J., and Escalona, M. (2017). Agile requirements engineering: a systematic literature review. *Comput. Standards Interfaces* 49, 79–91. doi: 10.1016/j.csi.2016.08.011
- Schwaber, K., and Sutherland, J. (2020). *The Scrum Guide*. Available online at: <https://www.scrumguides.org/scrum-guide.html> (accessed February 10, 2023).
- Sidky, A., Arthur, J., and Bohner, S. (2007). A disciplined approach to adopting agile practices: the agile adoption framework. *Innovat. Syst. Softw. Eng.* 3, 203–216. doi: 10.1007/s11334-007-0026-z
- Strode, D. E., Sharp, H., Barroca, L., Gregory, P., and Taylor, K. (2022). Tensions in organizations transforming to agility. *IEEE Trans. Eng. Manage.* 69, 3572–3583. doi: 10.1109/TEM.2022.3160415
- Susnjak, T. (2022). ChatGPT: The end of online exam integrity?. *arXiv[Preprint]*. arXiv: 2212.09292. doi: 10.48550/ARXIV.2212.09292
- The decoder. (2023). *Google launches ChatGPT competitor "Bard" and more*. Available online at: <https://the-decoder.com/google-launches-chatgpt-competitor-bard> (accessed February 9, 2023).
- Winkler, R., and Soellner, M. (2018). Unleashing the potential of chatbots in education: a state-of-the-art analysis. *Acad. Manage. Proc.* 2018, 15903. doi: 10.5465/AMBPP.2018.15903abstract
- Zawacki-Richter, O., Marin, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *Int. J. Educ. Technol. Higher Educ.* 16, 1–27. doi: 10.1186/s41239-019-0171-0



OPEN ACCESS

EDITED BY

Kamal Kant Hiran,
Symbiosis University of Applied Sciences, India

REVIEWED BY

Cucuk Wawan Budiyo,
Sebelas Maret University, Indonesia
Luis Pedro,
University of Aveiro, Portugal

*CORRESPONDENCE

Claudia Camacho-Zuñiga
✉ claudia.camacho@tec.mx

RECEIVED 30 June 2023

ACCEPTED 31 July 2023

PUBLISHED 17 August 2023

CITATION

Mondragon-Estrada E, Kirschning I,
Nolazco-Flores JA and
Camacho-Zuñiga C (2023) Fostering digital
transformation in education: technology
enhanced learning from professors'
experiences in emergency remote teaching.
Front. Educ. 8:1250461.
doi: 10.3389/feduc.2023.1250461

COPYRIGHT

© 2023 Mondragon-Estrada, Kirschning,
Nolazco-Flores and Camacho-Zuñiga. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Fostering digital transformation in education: technology enhanced learning from professors' experiences in emergency remote teaching

Enrique Mondragon-Estrada¹, Ingrid Kirschning²,
Juan Arturo Nolasco-Flores³ and Claudia Camacho-Zuñiga^{1,3*}

¹School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico, ²School of Engineering, Universidad de las Américas Puebla, Puebla, Mexico, ³Institute for the Future of Education, Tecnológico de Monterrey, Monterrey, Mexico

In this work, we aim to understand professors' perception of the key competences as well as the best educational strategies and technological tools to guide digital transformation (DT) in education, according to their experience in emergency remote teaching (ERT). In recent years, technological advancement has driven DT in many areas, with education being among them. ERT due to COVID-19 accelerated this transition. Restrictions and lockdowns forced higher education institutions to adopt remote teaching strategies and tools suited for a digital environment. We surveyed 100 professors from a private Mexican university with 15-month experience of online ERT. We asked them through Likert scale questions to self-evaluate their performance and whether they perceived it to be better in online or hybrid environments compared with face-to-face environments in different aspects. We performed correlation, cluster, and factor analysis to identify the relationships and patterns in their answers. Through open-ended questions, we also asked the participants about the challenges and achievements they experienced, and the educational strategies and technological tools they successfully incorporated during ERT. We also conducted text mining to extract the most relevant information from these answers and validated that they were not polarized with negative sentiment using a large language model. Our results showed social intelligence as an underlying competence for teaching performance was highlighted in the digital environment due to the physical interaction limitations. Participants found success in implementing information and communication technologies, resulting in maintaining student interest and building trust in the online environment. Professors recognized the relevance not only of learning management systems and communication platforms, as expected, but also hardware such as tablets, cameras, and headphones for the successful delivery of education in a digital environment. Technology Enhanced Learning transposes game-based, quizzing practices, and collaborative learning to digital environments. Furthermore, the professors recommended learning-by-doing, flipped learning, problem-based learning, game-based learning, and holistic education as some pedagogical methodologies that were successfully applied in ERT and could be implemented for DT. Understanding the gains concerning teaching learning strategies and technologies that were incorporated during ERT is of the utmost importance for driving DT and its benefits for current and future education.

KEYWORDS

information and communication technologies, higher education, technology enhanced learning, educational innovation, digital transformation in education, faculty experiences, emergency remote teaching

1. Introduction

The advent of digital computing has triggered a digital transformation (DT) in every human sector (Zaoui and Souissi, 2020). DT refers to the integration of information and communication technologies (ICT) as well as computing and connectivity technologies into an organization (Vial, 2019; Itten et al., 2020).

DT has introduced new mechanisms for how organizations can succeed, grow, and stay competitive (Kraus et al., 2022) by integrating IT infrastructure and digital resources with the appropriate organizational strategies (Zhang et al., 2023). It is about redesigning how an organization uses technological and human resources as well as processes to fundamentally change its performance. It is an essential factor of Industry 4.0 and is revolutionizing how humans address problems in different fields (Kraus et al., 2021). The purpose is to improve efficiency and productivity, customer expectations, data-driven decision making, innovation, and resilience and adaptability. Furthermore, DT also involves changes in leadership, different thinking, innovation and new business models, and the incorporation of digitization in every aspect of the organization (Klein, 2020).

Higher education institutions (HEIs) have the main purpose of preparing students to become professionals and part of the future workforce (Khan et al., 2022). The contemporary technological landscape with its fast and continuous evolution demands professionals be equipped with digital and technological skills (Akour and Alenezi, 2022). That is why HEIs should aim for DT to keep evolving along with companies and industries, and to fill the current labor market needs (Benavides et al., 2020). This is achieved in two ways: by fostering the progressive process of digital infrastructure from a bottom-up direction, where subjects are gradually digitalized; and top-down processes, where strategic and logistical needs are served by larger and shared solutions (Bygstad et al., 2022). Despite much work being done regarding DT in the business field, further work needs to be done in the higher education field.

DT in HEIs is attained through different strategies, including digital strategies. HEIs must incorporate the use of technological tools in the classroom (Abelha et al., 2020), such as delivering courses via distance education tools that consist of online synchronous transmission supported by computer tools such as video communication platforms (García-Morales et al., 2021), and with other support tools, such as the use of AI to detect the sentiment of the students and in this way measure the effectiveness of the lecture (Giang et al., 2020). A digital strategy also supports an operational optimization strategy by improving students' experience through providing technology to faculties, training on teaching methodologies, and tools to deliver these methodologies (Fernández et al., 2023). Finally, a digital strategy also supports a technological optimization strategy by selecting the best video communication platforms and working with the provider to maximize security issues, educational features, and quality of delivery (Alenezi, 2023). The challenge of

implementing these strategies is accompanied by the ability of faculties to adapt and use these tools in the best way. In this adaptation, faculties have to implement their own teaching strategies to be effective.

One of the largest leaps achieved toward DT in HEIs was forced by emergency and occurred during the COVID-19 pandemic. To stop the spread of the disease, governments around the world implemented lockdowns. All economic activities were affected, and they had to design and develop aggressive strategies to minimize the negative effects (Kumar et al., 2021). During this period, HEIs aimed to continue delivering high quality education and implemented several disruptive teaching and digital strategies. This is why the purpose of this work is to survey faculties to understand the challenges they faced and the strategies they found effective to deliver high quality courses while addressing DT. It was guided by the following research questions: (1) What teaching-learning strategies and technologies were successfully incorporated by teachers accustomed to face-to-face courses during remote education? and (2) What teaching-learning strategies and technologies do they recommend to other professors to effectively carry out online courses?

The current work will provide relevant information and insights on DT, supporting emergency remote teaching, to understand how DT can: ensure uninterrupted learning in the face of possible crises; increase access and equity to make education more accessible to students due to geographical, financial, or personal constraints; improve efficiency and innovation to achieve more efficient administrative processes and innovative teaching methods using digital tools; and ensure that graduates are fully equipped for the growing demand for digital skills.

2. Theoretical framework

Higher education institutions (HEIs) lead societal change and improvements. In this case, HEIs contribute to the digital development of the geographical regions where they operate. Teixeira and collaborators indicated that besides the acquisition of knowledge and skills, higher education contributes to job creation and to the improvement of the quality of life at work. They affirm that in recent years there has been a greater involvement between higher education and the labor market, namely, through research centers, business consultancy, and partnerships with local organizations (Teixeira et al., 2021).

DT inside HEIs goes beyond technological progress; it is a deep and transcendental evolution that generates changes in the mission, the philosophy, the pedagogical approaches, the teaching and learning processes, the research, the administrative activities, and even the relationships with other societal organizations. DT demands rethinking, restructuring, and reinventing HEIs from its multi-purpose, multi-processes, multidisciplinary, multistate, and

multi-stakeholders character. That is, DT should be an integral and holistic transformation of HEIs (Benavides et al., 2020).

From the educational dimension, DT involves integrating digital technologies into teaching, learning, and organizational practices as a resource to create additional and differentiated value for students and other stakeholders. DT is more than just digitizing information and processes; it is about leveraging technology to redesign educational processes, culture, and educational stakeholders' experiences to meet changing educational and labor market requirements (Benavides et al., 2020). Beyond students, the aim of DT is to benefit other educational stakeholders. These include: employers, parents, research funding agencies, and the community (clients); other universities and educational institutions, either local or international, formal, or informal, with physical or online campuses (competitors); and publishers, technology providers, faculty, staff, and other societal partners (providers).

The positive aspects of DT for administrative staff inside HEIs are manifested in improving information flow management, providing open access to educational resources and research results, and reducing the cost of higher education. Among the benefits of the DT of HEIs are ensuring a broader delivery of higher education through distance learning courses, thereby fostering its accessibility and mass scale, and making training and learning more flexible, personalized, and effective (Штыхно et al., 2020).

In this context, computer-supported collaborative learning (CSCL) is concerned with how information and communication technologies (ICTs) can support education (Ludvigsen and Mørch, 2010). The use of ICTs is an essential component for effective online teaching (Vikas and Mathur, 2022). New technologies have been used and applied to online learning, such as learning management systems (LMS), synchronization services, and social network tools (Bond et al., 2021).

Incorporating technology in education has been shown to have both negative and positive impacts throughout the years. On the one hand, it allows the non-academic use of technology which increases academic distraction (Dontre, 2021) and opens the possibility of academic dishonesty through cheating when searching for content on the internet (Golden and Kohlbeck, 2020), and excessive use of technology has been shown to be associated with lower academic performance (Gorjón and Osés, 2023). On the other hand, it enhances the learning experience (Dunn and Kennedy, 2019; Shen and Ho, 2020; Rosli and Saleh, 2022), improves accessibility (Coleman and Berge, 2018; Seale, 2020), fosters active learning (Theobald et al., 2020), promotes collaboration (Blau et al., 2020), and keeps pace with technological advancements (Miranda et al., 2021). Furthermore, it helps develop digital competence or digital literacy. This competence is defined as "the ability and disposition to use digital media, to develop them in a productive and creative way. It refers to the capacity to critically reflect on its usage and the impact media have on society and work, both for private and professional contexts, as well as the understanding of the potentials and limits of digital media and their effects" (Blau et al., 2020; Ehlers, 2021).

In response to students' expectations and the multiple benefits that DT provides, nowadays it is mandatory to foster DT inside HEIs. The use of ICT from a technological, informational, multimedia, communicative, collaborative, and ethical perspective is no longer optional for professors (Basilotta-Gómez-Pablos et al., 2022). They should strengthen their digital competence as an essential skill required for in the era of Industry 4.0 (Ehlers, 2021). Nevertheless, it has been discussed how educational institutions should provide training to teachers and professors in this matter to enable appropriate technology

pedagogy integration, since current efforts are not enough (Bhebhe et al., 2023; Pozas and Letzel, 2023).

In this regard, the COVID-19 pandemic was an opportunity to attain DT in education (Mhlanga et al., 2022), the outcomes of which will not only be helpful in health contingencies but also during other emergencies such as natural disasters or wars, for instance, the current war in Ukraine (Banyoi et al., 2023). The COVID-19 outbreak accelerated DT in HEIs through the pervasion of ICTs into the classrooms. It forced a migration from face-to-face courses to online and hybrid learning in a short period of time, which represented a challenge for educational institutions (Daniel, 2020). Even though online learning was already present in some HEIs around the world, many universities had to focus on the redesign of courses by implementing new strategies and technologies (Zhang et al., 2022) leading to sudden emergency remote teaching (ERT) practices (Hodges et al., 2020).

ERT forced every educational stakeholder to adapt to new virtual environments. Students' behavioral, cognitive, and affective engagement has been shown to be a relevant factor for online learning (Daher et al., 2021; Salas-Pilco et al., 2022; Yang et al., 2022). During distance learning, students themselves have acknowledged that strategies not usually applied in face-to-face courses have been necessary, such as concerning about their well-being and implementing additional technological tools (Utomo et al., 2020; George and Thomas, 2021; Mondragon-Estrada and Camacho-Zuniga, 2021). Moreover, online pedagogy requires teachers with skills and capabilities for implementing technological tools and new strategies (Llerena-Izquierdo and Ayala-Carabayo, 2021; Archambault et al., 2022). Educational institutions should assist teachers and professors through ICT support staff as well as adequate end-user training to avoid negative consequences in their well-being (Pozo-Rico et al., 2020; Wang and Zhao, 2023). For instance, technostress, defined as an adaptation problem to cope or get used to ICTs, is an undesired consequence of new computer technologies (Ahmad et al., 2012).

Professors and teachers must foster DT in the educational ecosystem. It has been suggested that by pedagogically pre-designing their courses using modern pedagogical technologies and preparing their lectures using ICTs, teachers make students learn those technological tools as a means of teaching (Olimov and Mamurova, 2022). This involves not only using some technological tools for some activities but also embracing a more flexible approach that could open new possibilities in the classroom (Feerick et al., 2022).

New technologies such as the metaverse (Misirlis and Munawar, 2023) and AI models such as ChatGPT (Mhlanga, 2023) have been proposed to be applied in education to further advance DT in HEIs. However, the fact remains that every time there are new technologies to be used requires teachers and professors to be properly prepared. Our study examines the opinions of professors that experienced ERT during COVID-19 lockdown and who advise educational strategies and technological tools that were, in their perspective, suitable for this challenge and those that will inevitably continue to exist in the growing technological age.

3. Materials and methods

3.1. Participants

In total, 100 faculty members from a private university in Mexico with experience teaching online courses for 15 months due to the

COVID-19 outbreak participated in this study voluntarily. They were informed about the research purpose of the survey and agreed to the anonymous and confidential use of their provided information.

From this sample, ($n=44$) were female, ($n=52$) male, and ($n=4$) unspecified. The professors belonged to different schools inside the university: ($n=22$) were from the School of Engineering, ($n=22$) from the School of Sciences, ($n=19$) from the School of Social Sciences, ($n=18$) from the School of Arts and Humanities, ($n=18$) from the School of Business and Economics, and ($n=1$) from the School of Research and Graduates. Their teaching experience was between 2 and 48 years (mean = 17.8 years, SD = 11.7 years).

3.2. Data collection

An online optional survey was disseminated among 289 full-time professors of a private university in Mexico via institutional email, from which 100 answered. The survey consisted of seven sections, the first one regarding sociodemographic information and the remaining six sections are explained below.

Section 1 consisted of seven Likert scale items (1 – Poor, 5 – Excellent) about the participants' perceptions of their performance during online or hybrid courses in different aspects: (1) motivation, trust, empathy, and ethical commitment (Empathy); (2) methodological teaching strategies (Methodol.); (3) use of information and communications technology (ICT); (4) communication with students (Comm.); (5) course content design (Design); (6) collaboration with other faculty members (Collab.); and (7) creating and applying new methodologies, resources, and knowledge (Innov.). We included these aspects inspired by the previously published literature on the competences that a professor or university teacher should possess (Perez-Poch and López, 2016) and added the use of ICT as a fundamental skill for the digital transformation of teaching.

Section 2 included seven items about whether the participants perceived their performance to be better in online and hybrid environments compared with face-to-face environments. They analyzed the same aspects evaluated in Section 1 as Likert scale questions (1 – The face-to-face model is much better, 5 – The hybrid or online model is much better).

Sections 3 and 4 corresponded to open-ended questions asking participants about their biggest challenges and achievements, respectively, as professors during online or hybrid courses. Finally, Sections 5 and 6 consisted of open-ended questions seeking participants' recommendations for educational strategies and technologies, respectively, to improve their teaching based on their experiences in online and hybrid courses.

3.3. Data processing and analysis

3.3.1. Investigating the relationships between competences

To find how aspects (Empathy, Methodol., ICT, Comm, Design, and Collab. Innov) related to each other, we performed correlation, cluster, and factor analysis for the Likert scale questions in Sections 1 and 2 using R (R Core Team, 2022). To find the extent to which the categorical variables moved in the same or opposite direction consistently, correlation matrices were computed using the non-parametric Spearman's rank

correlation coefficient along with their significance levels for each section. Since we were interested in knowing how similar these aspects are to each other and how they could be grouped, we performed hierarchical clustering via complete-linkage clustering using Gower's distance, a metric suitable for categorical data (Maechler, 2022). Factor loadings were found by performing maximum-likelihood factor analysis with different numbers of factors (one, two, or three) to be fitted for each section to find unobserved latent variables that explained the variability of the collected data. For the factor analysis, we assumed our sample size was large enough, so that the sampling distribution of the mean of any independent, random variable would be normal or nearly normal, according to the central limit theorem.

3.3.2. Faculty's strengths and weaknesses

For obtaining insights regarding professors' main challenges and achievements during online or hybrid courses, we were required to perform natural language text analysis on the responses to the open-ended questions of Section 3 and 4. This text analysis can be done through term frequency (TF) and term frequency – inverse document frequency (TF-IDF). TF-IDF is a statistical measure that evaluates how relevant a word is to a document in a collection of documents. Thus, TF-IDF was computed to extract key terms from faculty members' answers. It was calculated as follows (Silge and Robinson, 2016):

$$\text{TF-IDF}(t, q, Q) = \frac{f_q(t)}{\max_{w \in q} f_q(w)} \cdot \ln \left(\frac{|Q|}{|\{q \in Q : t \in q\}|} \right) \quad (1)$$

where $f_q(t)$ is the frequency of the term t in the answer to question q , w is the number of words in the answer to question q , and Q is the total number of questions.

3.3.3. Validating faculty members' polarity

Text sentiment analysis can be used to determine the sentiment and polarity of answers to challenges and achievements (Sections 3 and 4). Polarity depends on the balance of *objective fact statements* and *subjective non-fact statements* (Sahu and Majumdar, 2017). Facts are objective terms like events, entities, and their properties. On the other hand, a non-fact statement is subjective and usually related to an individual's sentiments, personal beliefs, opinion, perspective, feelings, or thoughts. Therefore, we performed sentiment analysis to confirm the objectivity of the faculty members' answers to the survey.

For this analysis, we used large language models (LLMs) because in the last few years these models have been shown to outperform classical data-driven models, such as Support Vector Machines (SVMs) and Recurrent Neural Networks (RNNs) for natural language processing (NLP) tasks (Fan et al., 2023). There are many LLMs but we used the RoBERTa (Pérez et al., 2021a,b) as our LLM, because it has been shown to outperform other state-of-the-art LLMs, such as RoBERTa, for sentimental analysis in Spanish.

3.3.4. Educational strategies and technological tools analysis

We were interested to know which educational strategies and technological tools were recommended to be applied before a lecture, during a lecture, and after a lecture. Therefore, for educational strategies, we carefully analyzed all answers and created a tree

structure to have a visual plot. We were also interested to know what type of technology is effective; therefore, we carefully analyzed all the faculty members' answers and created a matrix relating specific technology with categories. Additionally, we performed TF-IDF as described in Eq. 1 to contrast both results.

4. Results

4.1. Investigating the relationships between competences

Professors' performance involves multiple competences (Perez-Poch and López, 2016); Teaching is a complex activity where every factor that was analyzed in this work benefits the rest. Professors' self-evaluation (1 – Poor, 5 – Excellent) had a median value of 4 ($IQR=1$) for the seven aspects. Figure 1 depicts the heat map of the correlation matrix using Spearman's rank for the seven aspects that professors perceived about their own performance (dark blue = 1 and white = 0); the dendrogram resulting from hierarchical clustering is also shown. Every correlation is positive and significant ($p < 0.05$), reflecting that professors' performance involves every aspect evaluated in Section 1, and that each one of them contributes to the rest in a positively related manner.

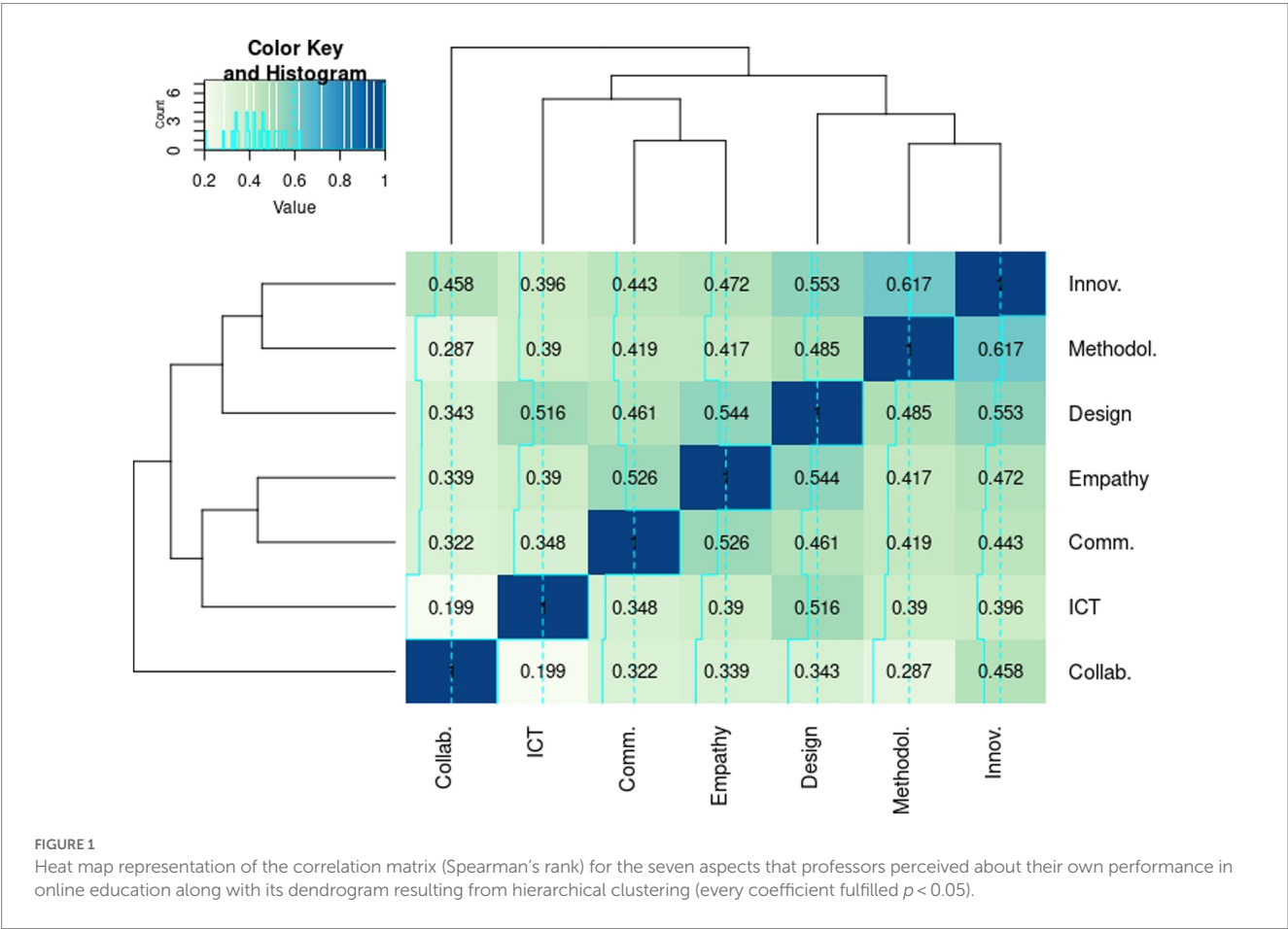
In the current technology-driven educational environment, professors recognized the relevance of innovation, that is, creating and applying new methodologies, resources, and knowledge, into teaching

methodology and course content design. Correlation analysis showed that, according to faculty members' perceptions of their own performance, the highest coefficients corresponded to methodology and innovation [$r(98) = 0.617$] together with innovation and course design [$r(98) = 0.553$]. This also highlights a unique opportunity for HEIs to improve, since in previous studies innovation was identified as inherent to online educational models (Camacho-Zuñiga et al., 2023).

In current research, collaborative work among faculty members showed little correlation with professors' own evaluation of their performance. The dendrogram from Figure 1 shows three relevant clusters:

- Innovation, methodology, and design: these concern the characteristics of the course;
- Communication, empathy, and ICTs: these pertain to the human interactions involved in the teaching and learning process and the means of communication; and
- Collaboration with other faculty members: this aspect was not merged into any other cluster and showed the lowest correlation with the other aspects.

Innovation can be understood as the combination of the four Cs of the 21st century skills: critical thinking, collaboration, creativity, and communication (Demirkol-Orak and İnözü, 2021). The fact that innovation was accompanied by methodology and design evidences



the importance of creativity and critical thinking applied to the course content design, which is in accordance with previous work (Demirkol-Orak and İnözü, 2021).

Collaborative work among faculty members has been reported to trigger professors' performance improvement (De Oca et al., 2021); in agreement, the current research revealed a positive correlation between innovation and faculty collaboration. Previous work (Garzón Artacho et al., 2020) has shown that use of ICTs has a direct relationship with communication, collaboration, and content creation. Similarly, in our analysis, communication and empathy were clustered with the use of ICTs, but collaboration was not clustered with the use of ICTs, but, in the opposite way, collaboration was clustered apart.

In Section 2, regarding the same seven aspects as in Section 1, professors expressed their opinion about the best teaching model in a Likert scale, where 1 represented face-to-face model, and 5 hybrid or online model. Face-to-face model was better at Empathy, with 32% of professors selecting 1. As could be expected, hybrid or online model was better at ICT, with 33% of professors selecting 4. Methodology, Communication, Design, Collaboration, and Innovation were equally evaluated in both models, with 28%–46% of professors selecting 3 for these aspects. Figure 2 displays a heat map corresponding to the Spearman's rank correlation coefficients of the professors' responses to Section 2 as well as the dendrogram from hierarchical clustering. All correlations are positive and significant ($p < 0.05$). The pairs with the highest correlation coefficients were course design–methodology,

teaching methodology–innovation, and course design–innovation. The three aspects are similar and were grouped in the same cluster, echoing the findings from Section 1 (see Figure 1). Moreover, communication and empathy are similar, and they, in turn, exhibited a similarity with both faculty collaboration and the use of ICTs.

The hierarchical clusterings reported in Figures 1, 2 suggest that the use of ICT in education either demands for strong faculty collaboration or that such collaboration facilitates the incorporation of ICTs and the ongoing innovation of online or hybrid learning.

Evaluations of teachers' performance usually consider teaching methodology and course content (Abbas et al., 2022); however, the current work evidences that the shift from face-to-face to online teaching highlighted social intelligence as a fundamental competence for professors' performance. Factor analysis was performed using one, two, and three factors along with a chi-square test (H_0 : n number of factors are sufficient to explain the set of observations). For faculty members' perceptions of their performance (Section 1), three factors were sufficient to explain the set of observations [$\chi^2(3, N = 100) = 0.74, p = 0.863$] Table 1 reports the loadings of those factors and the interpretation from the authors' point of view:

1. "Innovative teaching methodology": innovation (0.921), methodology (0.470), and faculty collaboration (0.390) have the highest loading values in this factor. The first two are related

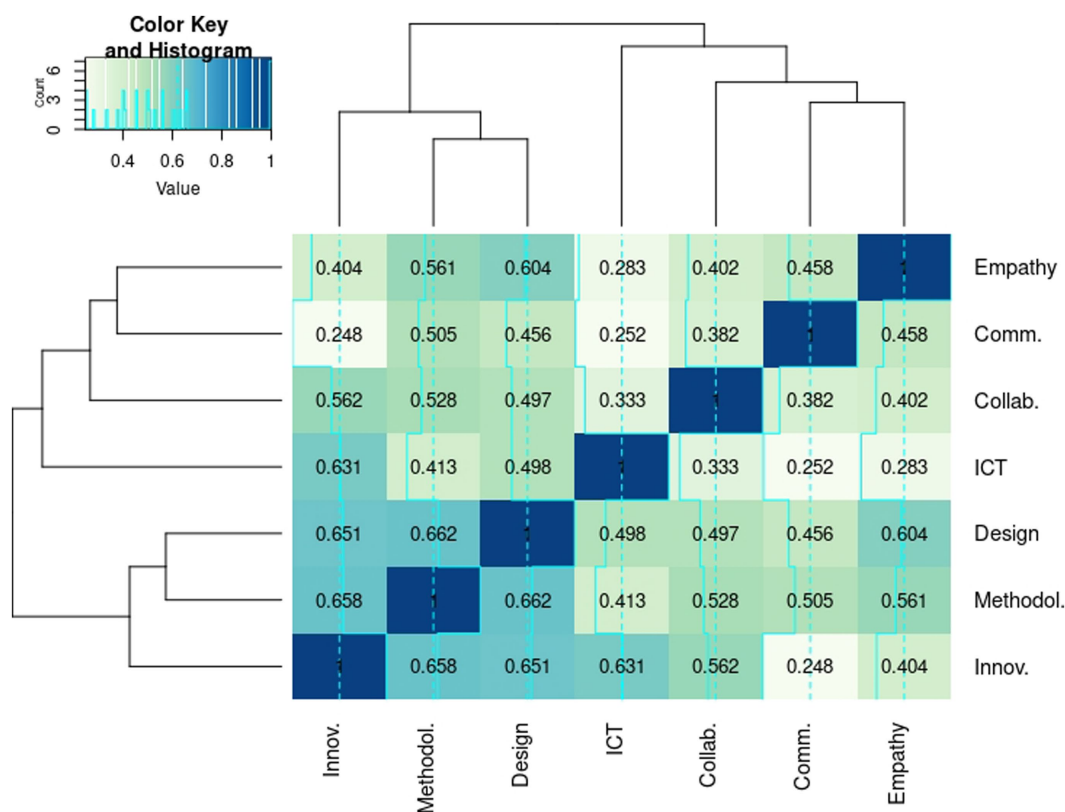


FIGURE 2

Heat map representation of the correlation matrix (Spearman's rank) and dendrogram resulting from hierarchical clustering for the seven aspects that professors ranked to be better between hybrid or online models and face-to-face models (every coefficient fulfilled $p < 0.05$).

TABLE 1 Loadings found for faculty members' perceptions of their own performance (Section 1 of the survey) during online education using factor analysis with three factors.

Category	Factor 1 "Innovative teaching methodology"	Factor 2 "Instructional design"	Factor 3 "Social intelligence and communication"
Empathy	0.242	0.373	0.592
Methodology	0.470	0.384	0.274
ICT	0.178	0.499	0.275
Communication	0.219	0.298	0.553
Design	0.288	0.702	0.273
Collaboration	0.390	0.108	0.319
Innovation	0.921	0.323	0.206

TABLE 2 Loadings found for faculty members' comparison of face-to-face vs. online education (Section 2 of the survey) during online education using factor analysis with three factors.

Category	Factor 1 "Technology-driven teaching methodology"	Factor 2 "Social intelligence and communication"	Factor 3 "Instructional design"
Empathy	0.240	0.598	0.315
Methodology	0.519	0.639	0.182
ICT	0.589	0.182	0.251
Communication	0.113	0.662	0.135
Design	0.456	0.455	0.762
Collaboration	0.496	0.418	0.161
Innovation	0.959	0.200	0.190

with creating and applying new resources, knowledge, and teaching methodologies, and the third pertains to collaboration with other faculty members to include ICTs and improving the teaching methodology.

2. "Instructional design": in this factor, design (0.702) holds the highest loading, followed by ICT (0.499) and empathy (0.373). Overall, it might suggest that empathizing with the student while designing the course material using ICTs is an underlying feature of course design.
3. "Social intelligence and communication": in this factor, empathy (0.592), communication (0.553), and faculty collaboration (0.319) have the highest loadings. All three pertain to human interactions, which might be an apt description of this latent variable.

For this research, we consider social intelligence as being able to generate effective environments for collaboration and negotiation in multicultural contexts with respect and appreciation for the diversity of people, knowledge, and cultures. Communication refers to effectively using different languages, resources, and strategies according to the context for an effective interaction within various professional and personal networks with different purposes or objectives (Instituto Tecnológico y de Estudios Superiores de Monterrey, 2018; Ehlers, 2021). Perez-Poch and López previously reported that working collaboratively with the rest of the faculty was perceived by professors as the least important teaching competence (Perez-Poch and López, 2016). In contrast, our evidence shows that this competence as a part of social intelligence is fundamental in both face-to-face and digital

supported interactions. Moreover, social intelligence and communication are valuable for interacting in heterogeneous groups in many areas of life through cooperation, participation, and motivation, which has been shown to be greatly beneficial for education for sustainable development, as one of the 17 sustainable development goals adopted by the United Nations (Scherak and Rieckmann, 2020).

Online education contrasts to face-to-face education by being technology-driven and highly innovative. Faculty perceptions of whether online and hybrid courses or face-to-face courses are better in several aspects were reduced to three factors [$\chi^2(3, N = 100) = 3.97, p = 0.265$]. Table 2 reports the loadings for each factor:

1. "Technology-driven teaching methodology": innovation (0.959), use of ICTs (0.589), and methodology (0.519) were the variables with highest loadings in this factor. These are closely tied to educational strategies and technology, suggesting that the application of new methodologies using ICTs is a key factor in educational transformation.
2. "Social intelligence and communication": this factor reflects the relevance of social intelligence competences in teaching performance, as the concepts with the highest loadings were communication (0.662), methodology (0.639), and empathy (0.598).
3. "Instructional design": in this factor, design surpassed the rest of the aspects (0.762), followed by empathy (0.315). It can be inferred that professors should consider fostering student motivation, trust, empathy, and ethical commitment while designing a course.

TABLE 3 Top terms with highest TF-IDF values for professors' greatest challenges during online education.

Term	N	TF	IDF	TF-IDF
Challenge (<i>reto</i>)	6	0.004	1.386	0.006
Capture (<i>captar</i>)	8	0.005	0.693	0.004
Really (<i>realmente</i>)	4	0.002	1.386	0.004
Attention (<i>atención</i>)	15	0.010	0.287	0.003
Sound (<i>sonido</i>)	3	0.002	1.386	0.003
Can (<i>poder</i>)	5	0.004	0.693	0.003
Feedback (<i>retroalimentación</i>)	5	0.004	0.693	0.003
Student (<i>estudiante</i>)	4	0.003	0.693	0.002
Interaction (<i>interacción</i>)	4	0.003	0.693	0.002

TABLE 4 Top terms with highest TF-IDF values for professors' greatest achievements during online remote education.

Term	N	TF	IDF	TF-IDF
Despite (<i>a pesar</i>)	8	0.006	1.386	0.008
New (<i>nuevas</i>)	8	0.006	0.693	0.004
Situation (<i>situación</i>)	4	0.003	1.386	0.004
Model (<i>modelo</i>)	6	0.004	0.693	0.003
Trust (<i>confianza</i>)	3	0.002	1.386	0.003
ITCs (<i>TICs</i>)	3	0.002	1.386	0.003
Achieve (<i>lograr</i>)	14	0.010	0.287	0.003
Distance (<i>distancia</i>)	5	0.004	0.693	0.002
Interest (<i>interés</i>)	10	0.007	0.287	0.002

The words "which," "day," and "what" were removed from the list since they do not provide additional information.

4.2. Faculty members' strengths and weaknesses

Professors' main challenges (Section 3) during online or hybrid courses were keeping students' attention and engagement, maintaining student-student and student-professor interactions, and providing truly efficient feedback. TF-IDF was performed on the text responses regarding professors' biggest challenges and the main results are reported in Table 3. As expected, "challenge" (0.006) and "student" were among the terms with highest TF-IDF values. Other relevant terms were "capture" (0.005), "attention" (0.003), "feedback" (0.003), and "interaction" (0.002). In agreement with Camacho-Zuñiga et al. (2023), it can be inferred that capturing students' attention, providing truly efficient feedback, and, in general, lecturing without the usual immediate and physical bidirectional communication of a face-to-face course were significant challenges for the professors.

Professors found success in implementing ICTs and redesigning their methodologies for the new teaching model, resulting in maintaining student interest and building trust in the online environment, despite the challenging situation imposed by the COVID-19 lockdown. The terms with the highest TF-IDF values for professors' biggest achievements are revealed in Table 4. The presence of "despite", "new", "situation", and "distance" are understandable as means to describe the challenging environment for teaching. Other terms in this chart are "trust", "new", "model", and "ICTs", the last two with the highest TF-IDF values. The term

'trust' could refer to building trust with students in the online environment or gaining confidence in using digital tools; its relatively high IDF score indicates this was a significant achievement, but not as common. The presence of "ICTs" suggests professors made significant strides in implementing and using technology in their teaching which could also benefit the ability of the professors to maintain or stimulate the interest of the students in online education.

4.3. Validating faculty members' polarity

Sentiment analysis of the respondents' answers to their *biggest challenge* and their *greatest achievement* while shifting to online learning was mostly neutral, which might be evidence of the objectivity of their answers. Sentiment analysis categorizes each answer to be positive, neutral, and negative; The sum of the predicted probabilities for each category is one. Sentiment analysis for faculty members' answers to their *biggest challenges* resulted in 31 positive, 53 neutral, and 16 negative answers. Below, we report some examples of positive, neutral, and negative classified answers as well as their sentiment score. It is noteworthy that most of the negative classified answers related to the fact that professors and students were not sharing the same physical space and, therefore, professors were unable to get immediate feedback from students' reactions.

"I've actually gotten along well." (*En realidad me he acoplado bien.*) Sentiment score predicted: Positive: 0.780, Neutral: 0.201, Negative: 0.019.

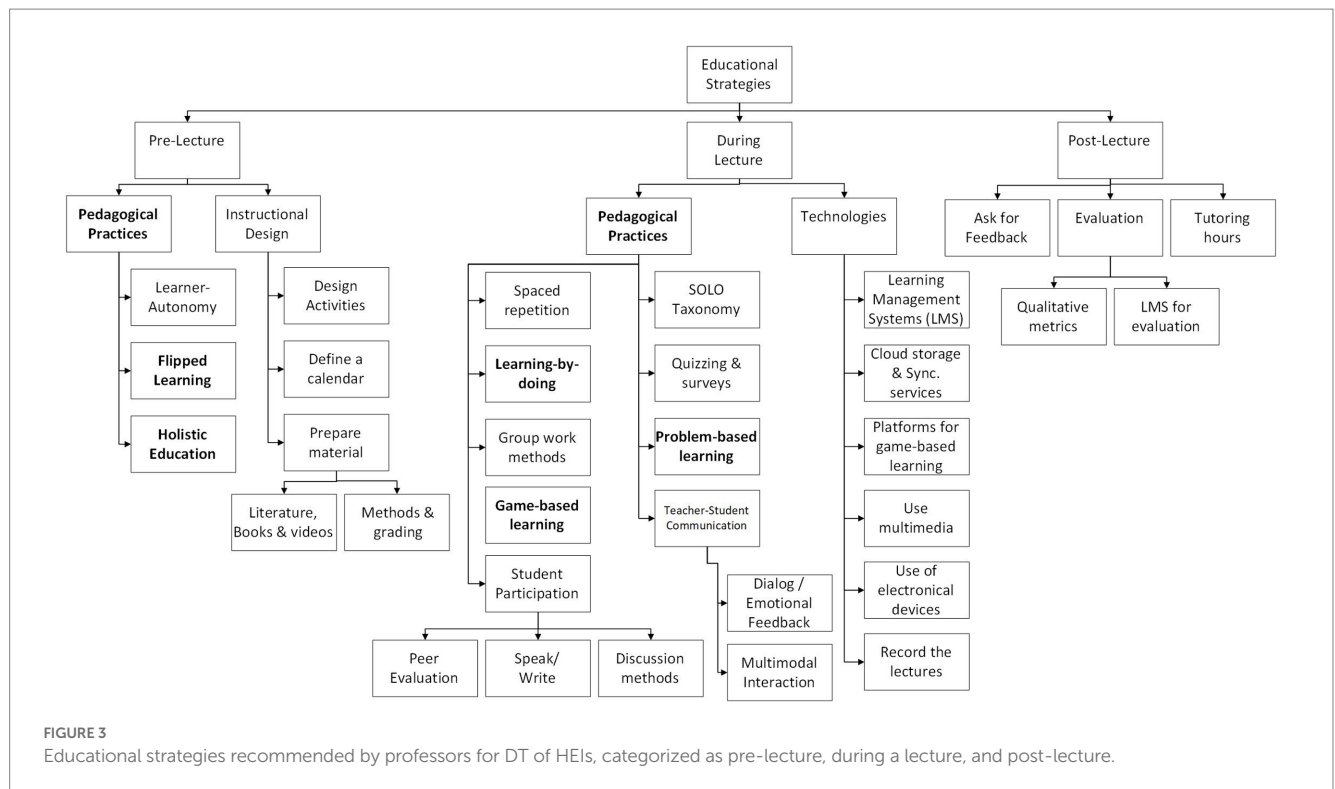
"Only the initial adaptation period. Most of the content was already prepared in digital format." (*Solo el periodo de adaptacion inicial. La mayoría de los contenidos ya los tenía preparados en formato digital.*) Neutral: 0.906, Positive: 0.052, Negative: 0.042.

"Technical failures and not being able to see students during the session; thus, it has been very difficult to have non-verbal feedback of what is discussed." (*Las fallas técnicas y el no poder ver a los estudiantes durante la sesión; así ha sido muy difícil tener una retroalimentación no verbal de lo que se discute.*) Negative: 0.980, Neutral: 0.018, Positive: 0.002.

Whereas the output for faculty members' answers to their *greatest achievements* were 46 positive, 52 neutral, and 1 negative. This shows that most of the answers were not polarized for negative sentiment. Examples of positive, neutral, and negative classified answers (and their sentiment scores) for this question are:

"Show to students that despite the limitations of the pandemic, interesting and high-quality projects can be carried out." (*Demostrar a los estudiantes que a pesar de las limitantes propias de la pandemia se pueden hacer proyectos interesantes y de calidad.*) Sentiment score predicted: Positive: 0.865, Neutral: 0.116, Negative: 0.019.

"Definitely the use of ICTs (platforms such as Blackboard, Teams, etc.)." (*Definitivamente el uso de las TICs (plataformas como Blackboard, Teams, etc.).*) Neutral: 0.829, Positive: 0.109, Negative: 0.062



“Provide a critical and participatory space in which students do not lose interest despite how complex and tiring it is for them to spend so many hours in front of the screen in their individual spaces.” (Generar un espacio crítico y participativo en el que los estudiantes no pierdan el interés a pesar de lo complejo y pesado que es para ellos estar tantas horas frente a la pantalla en sus espacios individuales.)
Negative: 0.692, Neutral: 0.261, Positive: 0.047

Overall, most professors highlighted the use of ICTs as their greatest achievement, as well as the fact that they captured students’ attention through innovative activities and by using technological tools.

4.4. Educational strategies and technological tools analysis

Multiple educational strategies were recommended by professors for an online or hybrid course (Section 5); however, the implementation of activities promoting interaction, collaborative learning, and learner autonomy were highlighted. Strategies were divided into three categories based on the time they could be applied: before a lecture, during a lecture, and after a lecture. Furthermore, during-lecture recommendations were divided into pedagogical and technological (see Figure 3).

During the lecture, technologies recommended comprise learning management systems (LMS) for organizing the content and grading works, cloud storage and synchronization services for collaborative activities, game-based learning platforms for interactive learning, use of multimedia such as videos and pictures, use of electronic devices, and recording the lectures for future reference. Professors also recommended the implementation of spaced repetition, learning-by-doing, and group work methods, as well as enhancing students’

participation during the lecture through the technologies previously mentioned.

Strategies recommended before a lecture are related to the design of the course material, calendar, and activities. Learner autonomy, flipped learning, and holistic education are some pedagogical methodologies that were recommended by faculty members and that could be applied during the design of a course. On the other hand, after-lecture recommendations included grading and evaluation systems, providing tutoring hours, and asking the students for feedback on how they are understanding the topics.

Some of the key terms resulting from TF-IDF analysis of the professors’ recommendations concerning educational methodologies were “content” (0.003), “collaborative” (0.002), “teams” (0.002), “course” (0.002), “flipped” (0.002), and “interactive” (0.002). These highlight the importance of the course content and working in teams. Also, the term ‘flipped’ appeared, pointing out the instructional strategy of flipped learning.

According to respondents’ recommendations, the pedagogical strategies most successfully implemented in the online environment and that, therefore, could enhance DT in HEIs are:

- **Learning-by-doing:** a principle that states how people learn from experiences and actions, rather than by observing or listening to the instructor (Reese, 2011). This practical experience helps students to apply their knowledge easily to real life situations (Bradberry and De Maio, 2019). Furthermore, it has been shown that learning-by-doing plays a key role in digital literacy since students learn how to use technology by using it (Tan and Kim, 2015).
- **Problem-based learning:** this is a pedagogical approach that enables students to learn while engaging actively with meaningful problems (Yew and Goh, 2016). The process starts with a case or problem statement which students, through a systematic

approach, work on addressing. Electronic environments provide an ideal environment to enhance problem-based learning through lab simulators, video presentations, and graphic tools (Moust et al., 2021).

- Game-based learning: an educational approach for engaging appropriate cognitive processes by playing games, participating in game-related communities, or by using worked examples in the games (Tobias et al., 2014). Game-based learning has evolved over the years, from digital games, online applications, and videos to emerging technologies such as virtual and augmented reality (Yu et al., 2022).
- Flipped learning: this is a teaching approach in which the conventional notion of classroom-based learning is inverted. Students are introduced to the learning material before class and problem-solving activities along with discussion with peers is carried out in the classroom (El Miedany, 2019). Additionally, flipped learning has been empowered by the use of ICTs, which encourages autonomous work, motivates students, helps the analysis of the content, and facilitates the communication between students and teachers (Cueva and Inga, 2022).
- Holistic education: this refers to the educational vision that cares for the human being as a whole, which encapsulates balance, inclusion, and connectedness as principles (Miller, 2019). Since it relates to people's mindset directly, it can be greatly beneficial to enhance personal and social transformation to tackle current world problems (Miseliunaite et al., 2022).

Concerning technological implementations for online education, in agreement with previous literature, the professors recognized the relevance of LMS and communication platforms; even more, current research demonstrates that hardware, like Wacom tablets, digital pencils (Peimani and Kamalipour, 2021), microphones, headphones, and lightning accessories (Serrianni and Coy, 2014), are highly recommended for online teaching. Professors' recommendations regarding technological tools for online or hybrid teaching were extracted from the text answers of Section 6. They were classified into different categories based on the nature or use of the tool (see Table 5). The first two categories are keys for online courses: LMS and communication platforms provide tools for managing course content and students' grades, as well as provide means of communication between professors and students. Furthermore, hardware was the third most repeated category, which is evidently necessary for efficient online teaching, including cameras, tablets, Wacom tablets, smartphones, and earphones. It is also noteworthy the relevance of cloud storage and synchronization systems, like Google Sites, that enable and promote collaborative learning.

Our results evidence that TEL spreads game-based and quizzing practices to online environments. Although quizzing and game-based learning are usual strategies in a face-to-face classroom, the participants recommended game-based platforms and Q&A/polling platforms—such as Kahoot!, Nearpod, Quizizz, Quizlet, Mentimeter, and Tweedback—as tools for engaging students during the online class. This agrees with other studies reporting that game-based learning fosters learning motivation in students (Wati and Yuniawatika, 2020; Sonsona et al., 2021; Krouska et al., 2022).

The biggest challenges reported during ERT were related to experimental activities and the strengthening of motor skills inside STEM (Science, Technology, Engineering, and Mathematics)

laboratories, which in the digital world were addressed through the use of mathematical software and simulators. For Section 6, TF-IDF was also computed. Some of the terms with the highest TF-IDF values were technological tools such as (Microsoft) Teams (0.007), YouTube (0.003), Padlet (0.003), and Google (0.003). Interesting concepts such as “platform” (0.005), “videos” (0.005), and “simulators” (0.004) also appeared. Simulators are related to engineering courses like biology, chemistry, mechanics, and electronics. This agrees with previous literature where professors have used simulators for their STEM and clinical courses (Tabatabai, 2020; Mamani et al., 2021; Pradhan and Madihally, 2022).

The current study has contributed to understanding the pedagogical practices and technologies that drive digital transformation successfully. This is not just beneficial but essential for HEIs; however, we must recognize some limitations concerning this research. It was conducted in a private Mexican university, which might not be fully representative of other HEIs due to the digital gap that characterizes the Latin America region (Galperin and UNESCO Office Montevideo and Regional Bureau for Science in Latin America and the Caribbean, 2017). Furthermore, this study was conceived during the COVID-19 pandemic, long before the release of ChatGPT into the public domain. Since then, a lot of AI tools have gained popularity and will impact higher education and its DT. These technologies could be addressed in a future study.

5. Conclusion

Advances in technology are fast paced, faster than our ability to incorporate its benefits into education; COVID-19 lockdown helped us to accelerate digital transition, but we are still far from Digital Transformation (DT) and its benefits. That is why knowing which teaching-learning strategies and technologies have been successfully incorporated by professors during Emergency Remote Teaching (ERT) is of the utmost importance for extrapolating them to current and future education.

This work collected the perceptions and recommendations of teaching-learning strategies and educational tools from 100 faculty members from a private university in Mexico, with an average teaching experience of 17.8 years and who had been delivering online courses for 15 months due to COVID-19 lockdowns. Our findings show that under the current technology-driven educational environment and toward a DT of higher education institutions (HEIs):

1. Innovation is a fundamental feature for better course content design, that is, creating and applying new methodologies.
2. Social intelligence and communication are underlying competences for professors' performance that were highlighted in the digital environment due to interaction limitations in this context.
3. The pedagogies considered to be most conducive for DT are learning-by-doing, problem-based learning, game-based learning, flipped learning, and holistic education. Professors successfully applied them in ERT and, furthermore, they could be implemented in online, face-to-face, or hybrid courses.
4. Professors found success in implementing information and communication technologies (ICTs) and redesigning their methodologies for the new teaching model, resulting in

TABLE 5 Technological tools recommended by participants for the DT of HEIs, and carefully classified by the authors in different categories based on the nature or use of the tool.

Technological Tool	Learning management system (LMS)	Communication platform /instant messaging	Hardware	Cloud storage / synchronization service	Presentation software	Game-based learning platform	Software for mathematics and simulation	Q&A / polling platform	Virtual white board	Videoconferencing tool	Online video platforms (OVP) / streaming	Screen casting	Note-taking software	Quiz platform	Online social annotation platform
Frequency	37	15	12	11	6	6	6	5	5	3	3	3	2	1	1
Blackboard	x	x		x										x	
Canvas	x	x		x										x	
Zoom		x						x	x	x	x	x			x
Teams		x		x				x		x	x	x		x	
Microsoft whiteboard				x					x				x		x
Microsoft office suite		x		x	x								x		
Google sites		x		x	x			x					x	x	
Miro				x	x								x		
Padlet				x	x								x		
Mentimeter				x	x			x							x
Genially				x	x										
Tweedback								x							x
Awingu, VPN software															
Idroo		x		x	x				x				x		
Blogger															
Monkey survey				x				x						x	
Kahoot!						x		x						x	
Nearpod					x	x		x						x	x
Quizizz						x								x	
Quizlet						x								x	x
Perusall								x							x
Slido		x			x			x							x
Backchannel chat		x						x							x
Mural				x					x				x		
YouTube										x	x				
OBS studio											x	x			
Mathematica							x								
GeoGebra							x								
Camera			x												
Wacom and other tablets			x												
Monitors			x												
Phones and smartphones			x												
Earphones			x												

maintaining student interest and building trust in the online environment.

5. Professors recognized the relevance not only of learning management systems (LMS) and communication platforms, as could be expected, but also hardware such as camera, earphones, and Wacom and other tablets that facilitate the successful delivery of a digital lecture.
6. Technology Enhanced Learning (TEL) spreads game-based, quizzing practices, and collaborative learning to the online environments, while simulators cover those course contents that used to take place in STEM laboratories.

Our research evidences the importance of incorporating TEL and ICTs in education from the classroom to advance the DT of HEIs. To fulfill the current industry requirements, it is important that these practices not only remain in a remote education context but also become an essential part in present and future education. Technology and science evolve rapidly, and it is the duty of professors and HEIs to prepare students, the next decision-makers, to be properly equipped to address the challenges and opportunities that might arise while shaping the future of the world.

Data availability statement

The datasets presented in this article are not readily available because the informed consent was provided by participants under the condition of anonymity and confidentiality of their answers. Requests to access the datasets should be directed to CC-Z, claudia.camacho@tec.mx.

Ethics statement

Ethical approval was not required for the studies involving humans because no personal or identifiable information was provided by the participants and ethical approval is not required under the local

legislation of the institutions that took part in this research. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CC-Z, JN-F, and IK contributed to the conception and design of the study. IK collected the data. EM-E conducted the data processing and analysis. EM-E, JN-F, and CC-Z wrote the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to acknowledge the support of the Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico, in the production of this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Abbas, A., Haruna, H., Arrona-Palacios, A., Camacho-Zuñiga, C., Núñez-Daruich, S., Enriquez de la O, J. F., et al. (2022). Students' evaluations of teachers and recommendation based on course structure or teaching approaches: an empirical study based on the institutional dataset of student opinion survey. *Educ. Inf. Technol.* 27, 12049–12064. doi: 10.1007/s10639-022-11119-z
- Abelha, M., Fernandes, S., Mesquita, D., Seabra, F., and Ferreira-Oliveira, A. T. (2020). Graduate employability and competence development in higher education—a systematic literature review using PRISMA. *Sustainability* 12:5900. doi: 10.3390/su12155900
- Ahmad, U. N. U., Amin, S. M., and Ismail, W. K. W. (2012). The relationship between technostress creators and Organisational commitment among academic librarians. *Procedia. Soc. Behav. Sci.* 40, 182–186. doi: 10.1016/j.sbspro.2012.03.179
- Akour, M., and Alenezi, M. (2022). Higher education future in the era of digital transformation. *Educ. Sci.* 12:784. doi: 10.3390/educsci12110784
- Alenezi, M. (2023). Digital learning and digital institution in higher education. *Educ. Sci.* 13:88. doi: 10.3390/educsci13010088
- Archambault, L., Leary, H., and Rice, K. (2022). Pillars of online pedagogy: a framework for teaching in online learning environments. *Educ. Psychol.* 57, 178–191. doi: 10.1080/00461520.2022.2051513
- BANYOI, V., KHARKIVSKA, O., SHKURKO, H., and YATSKIV, M. (2023). Tools for implementing distance learning during the war: experience of Uzhhorod National University, Ukraine. *Arab World Eng. J.* 1, 75–87. doi: 10.24093/awej/comm1.6
- Basilotta-Gómez-Pablos, V., Matarranz, M., Casado-Aranda, L. A., and Otto, A. (2022). Teachers' digital competencies in higher education: a systematic literature review. *Int. J. Educ. Technol. High. Educ.* 19:8. doi: 10.1186/s41239-021-00312-8
- Benavides, L., Tamayo Arias, J., Arango Serna, M., Branch Bedoya, J., and Burgos, D. (2020). Digital transformation in higher education institutions: a systematic literature review. *Sensors* 20:3291. doi: 10.3390/s20113291
- Bhebe, S., Schlebusch, L., and Gawie, S. (2023). Sustaining the integration of technology pedagogies in higher education after the COVID-19 pandemic. *Int. J. Learn. Teach. Educ. Res.* 22, 1–19. doi: 10.26803/ijlter.22.2.1
- Blau, I., Shamir-Inbal, T., and Avdiel, O. (2020). How does the pedagogical design of a technology-enhanced collaborative academic course promote digital literacies, self-regulation, and perceived learning of students? *Internet High. Educ.* 45:100722. doi: 10.1016/j.iheduc.2019.100722
- Bond, M., Bedenlier, S., Marín, V. I., and Händel, M. (2021). Emergency remote teaching in higher education: mapping the first global online semester. *Int. J. Educ. Technol. High. Educ.* 18:50. doi: 10.1186/s41239-021-00282-x
- Bradberry, L. A., and De Maio, J. (2019). Learning by doing: the long-term impact of experiential learning programs on student success. *J. Polit. Sci. Educ.* 15, 94–111. doi: 10.1080/15512169.2018.1485571
- Bygstad, B., Øvrelid, E., Ludvigsen, S., and Dæhlen, M. (2022). From dual digitalization to digital learning space: exploring the digital transformation of higher education. *Comput. Educ.* 182:104463. doi: 10.1016/j.compedu.2022.104463

- Camacho-Zuñiga, C., Peña-Becerril, M., Cuevas-Cancino, M. O., and Avilés-Rabanales, E. G. (2023). Gains from the transition from face-to-face to digital/online modality to improve education in emergency. *Front. Educ.* 8:1197396. doi: 10.3389/feduc.2023.1197396
- Coleman, M., and Berge, Z. L. (2018). A review of accessibility in online higher education. *Online J. Dist. Learn. Adm.* 21, 1–7.
- Cueva, A., and Inga, E. (2022). Information and communication Technologies for Education Considering the flipped learning model. *Educ. Sci.* 12:207. doi: 10.3390/educsci12030207
- Daher, W., Sabbah, K., and Abuzant, M. (2021). Affective engagement of higher education students in an online course. *Emerg. Sci. J.* 5, 545–558. doi: 10.28991/esj-2021-01296
- Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects* 49, 91–96. doi: 10.1007/s11125-020-09464-3
- De Oca, S. M., Villada-Balbuena, M., and Camacho-Zuniga, C. (2021). “Professors’ concerns after the shift from face-to-face to online teaching amid COVID-19 contingency: an educational data mining analysis” in *2021 machine learning-driven digital Technologies for Educational Innovation Workshop* (Monterrey, Mexico: IEEE), 1–5. doi: 10.1109/IEEECONF53024.2021.9733778
- Demirkol-Orak, S., and İnöz, J. (2021). Teachers awareness and actual practices of 21st century learning and innovation skills. *Int. Online J. Educ. Teach.* 8, 975–997.
- Dontre, A. J. (2021). The influence of technology on academic distraction: a review. *Hum. Behav. Emerg. Technol.* 3, 379–390. doi: 10.1002/hbe2.229
- Dunn, T. J., and Kennedy, M. (2019). Technology enhanced learning in higher education: motivations, engagement and academic achievement. *Comput. Educ.* 137, 104–113. doi: 10.1016/j.compedu.2019.04.004
- Ehlers, U.-D. (2021). *Future skills: the future of learning and higher education*. Norderstedt: BoD – Books on Demand.
- El Miedany, Y. (2019). “Flipped learning” in *Rheumatology teaching* (Cham: Springer International Publishing), 285–303. doi: 10.1007/978-3-319-98213-7_15
- Fan, L., Li, L., Ma, Z., Lee, S., Yu, H., and Hemphill, L. (2023). A bibliometric review of large language models research from 2017 to 2023. *arXiv* 2304.02020. doi: 10.48550/ARXIV.2304.02020
- Feerick, E., Clerkin, A., and Cosgrove, J. (2022). Teachers’ understanding of the concept of “embedding” digital technology in education. *Irish Educ. Stud.* 41, 27–39. doi: 10.1080/03323315.2021.2022521
- Fernández, A., Gómez, B., Binjaku, K., and Meçe, E. K. (2023). ‘Digital transformation initiatives in higher education institutions: a multivocal literature review’. *Educ. Inf. Technol.* 1–32. doi: 10.1007/s10639-022-11544-0
- Galperin, H. UNESCO Office Montevideo and Regional Bureau for Science in Latin America and the Caribbean (2017) ‘Digital society: gaps and challenges for digital inclusion in Latin America and the Caribbean’. UNESCO Office Montevideo and Regional Bureau for Science in Latin America and the Caribbean. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000262860_eng.
- García-Morales, V. J., Garrido-Moreno, A., and Martín-Rojas, R. (2021). The transformation of higher education during the COVID disruption: emerging challenges in an online learning scenario. *Front. Psychol.* 12:616059. doi: 10.3389/fpsyg.2021.616059
- Garzón Artacho, E., Martínez, T. S., Ortega Martín, J. L., Marín Marín, J. A., and Gómez García, G. (2020). Teacher training in lifelong learning—the importance of digital competence in the encouragement of teaching innovation. *Sustainability* 12:2852. doi: 10.3390/su12072852
- George, G., and Thomas, M. R. (2021). Quarantined effects and strategies of college students – COVID-19. *Asian Educ. Dev. Stud.* 10, 565–573. doi: 10.1108/AEDS-04-2020-0054
- Giang, N. T. P., Dien, T. T., and Khoa, T. T. M. (2020). “Sentiment analysis for university students’ feedback” in *Advances in information and communication*. eds. K. Arai, S. Kapoor and R. Bhatia (Cham: Springer International Publishing (Advances in Intelligent Systems and Computing)), 55–66. doi: 10.1007/978-3-030-39442-4_5
- Golden, J., and Kohlbeck, M. (2020). Addressing cheating when using test bank questions in online classes. *J. Account. Educ.* 52:100671. doi: 10.1016/j.jaccedu.2020.100671
- Gorjón, L., and Osés, A. (2023). The negative impact of information and communication technologies overuse on student performance: evidence from OECD countries. *J. Educ. Comput. Res.* 61, 723–765. doi: 10.1177/07356331221133408
- Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., and Bond, M. A. (2020). “The difference between emergency remote teaching and online learning”. 27. Available at: <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>.
- Instituto Tecnológico y de Estudios Superiores de Monterrey (2018) ‘TEC 21 educational model’. Instituto Tecnológico y de Estudios Superiores de Monterrey. Available at: <https://internationalfaculty.tec.mx/sites/g/files/vgjoivo1851/files/tec21-model.pdf>.
- Itten, R., Hischier, R., Andrae, A. S. G., Bieser, J. C. T., Cabernard, L., Falke, A., et al. (2020). Digital transformation—life cycle assessment of digital services, multifunctional devices and cloud computing. *Int. J. Life Cycle Assess.* 25, 2093–2098. doi: 10.1007/s11367-020-01801-0
- Khan, N., Sarwar, A., Chen, T. B., and Khan, S. (2022). Connecting digital literacy in higher education to the 21st century workforce. *Knowl. Manag. E-Learn. Int. J.* 14, 46–61. doi: 10.34105/j.kmel.2022.14.004
- Klein, M. (2020). Leadership characteristics in the era of digital transformation. *Bus. Manag. Stud. Int. J.* 8, 883–902. doi: 10.15295/bmij.v8i1.1441
- Kraus, S., Durst, S., Ferreira, J. J., Veiga, P., Kailer, N., and Weinmann, A. (2022). Digital transformation in business and management research: an overview of the current status quo. *Int. J. Inf. Manag.* 63:102466. doi: 10.1016/j.ijinfomgt.2021.102466
- Kraus, S., Jones, P., Kailer, N., Weinmann, A., Chaparro-Banegas, N., and Roig-Tierno, N. (2021). Digital transformation: an overview of the current state of the art of research. *SAGE Open* 11:215824402110475. doi: 10.1177/21582440211047576
- Krouska, A., Troussas, C., and Sgouroupolou, C. (2022). Mobile game-based learning as a solution in COVID-19 era: modeling the pedagogical affordance and student interactions. *Educ. Inf. Technol.* 27, 229–241. doi: 10.1007/s10639-021-10672-3
- Kumar, V., Alshazly, H., Idris, S. A., and Bourouis, S. (2021). Evaluating the impact of COVID-19 on society, environment, economy, and education. *Sustainability* 13:13642. doi: 10.3390/su132413642
- Llerena-Izquierdo, J., and Ayala-Carabayo, R. (2021). “University teacher training during the COVID-19 emergency: the role of online teaching-learning tools” in *Information technology and systems*. eds. A. Rocha et al. (Cham: Springer International Publishing (Advances in Intelligent Systems and Computing)), 90–99.
- Ludvigsen, S. R., and Mørch, A. I. (2010). Computer-supported collaborative learning: basic concepts, multiple perspectives, and emerging trends. *Int. Encycl. Educ.* Elsevier 5, 290–296. doi: 10.1016/B978-0-08-044894-7.00493-0
- Maechler, M. (2022) ‘Cluster: cluster analysis basics and extensions’. Available at: <https://CRAN.R-project.org/package=cluster>.
- Mamani, N. M., García-Peñalvo, F. J., Conde, M. Á., and Gonçalves, J. (2021) ‘A systematic mapping about simulators and remote laboratories using hardware in the loop and robotic: developing STEM/STEAM skills in pre-university education’, In 2021 international symposium on computers in education (SIIE). 2021 international symposium on computers in education (SIIE), Malaga, Spain: IEEE, pp. 1–6. doi: 10.1109/SIIE53363.2021.9583622
- Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *SSRN Electr. J.* 1–19. doi: 10.2139/ssrn.4354422
- Mhlanga, D., Denhere, V., and Moloi, T. (2022). COVID-19 and the key digital transformation lessons for higher education institutions in South Africa. *Educ. Sci.* 12:464. doi: 10.3390/educsci12070464
- Miller, J. P. (2019) *The holistic curriculum. Third edition*. Toronto Buffalo London: University of Toronto Press.
- Miranda, J., Navarrete, C., Noguez, J., Molina-Espinosa, J. M., Ramírez-Montoya, M. S., Navarro-Tuch, S. A., et al. (2021). The core components of education 4.0 in higher education: three case studies in engineering education. *Comput. Electr. Eng.* 93:107278. doi: 10.1016/j.compeleceng.2021.107278
- Miselunaitė, B., Kliziene, I., and Cibulskas, G. (2022). Can holistic education solve the World’s problems: a systematic literature review. *Sustainability* 14:9737. doi: 10.3390/su14159737
- Misirlis, N., and Munawar, H. B. (2023). An analysis of the technology acceptance model in understanding university students behavioral intention to use metaverse technologies. *arXiv* 2302.02176. doi: 10.48550/ARXIV.2302.02176
- Mondragon-Estrada, E., and Camacho-Zuniga, C. (2021). “Undergraduate’s perspective on being an effective online student during lockdown due to COVID-19 pandemic: an educational data mining study” in *2021 machine learning-driven digital Technologies for Educational Innovation Workshop* (Monterrey, Mexico: IEEE), 1–5. doi: 10.1109/IEEECONF53024.2021.9733773
- Moust, J., Bouhuijs, P., and Schmidt, H. (2021). *Introduction to problem-based learning: a guide for students*. 4th Edn, London: Routledge.
- Olimov, S. S., and Mamurova, D. I. (2022). Information Technology in Education. *J. Adv. Res. Sci. Prog.* 1, 17–22.
- Peimani, N., and Kamalipour, H. (2021). Online education and the COVID-19 outbreak: a case study of online teaching during lockdown. *Educ. Sci.* 11:72. doi: 10.3390/educsci11020072
- Pérez, J. M., Furman, D. A., Alemany, L. A., and Luque, F. (2021a). RoBERTuito: a pre-trained language model for social media text in Spanish. *arXiv* 2111.09453. doi: 10.48550/ARXIV.2111.09453
- Pérez, J. M., Giudici, J. C., and Luque, F. (2021b). Psysentimiento: a Python toolkit for sentiment analysis and SocialNLP tasks. *arXiv* 2106.09462. doi: 10.48550/ARXIV.2106.09462
- Perez-Poch, A., and López, D. (2016) ‘Do differences exist between how engineering and non-engineering lecturers perceive the importance of teaching competences?’, in 2016 IEEE Frontiers in Education Conference (FIE), Erie, PA, USA: IEEE, pp. 1–8.

- Pozas, M., and Letzel, V. (2023). "Do you think you have what it takes?" – exploring predictors of pre-service teachers Prospective ICT Use. *Technol. Knowl. Learn.* 28, 823–841. doi: 10.1007/s10758-021-09551-0
- Pozo-Rico, T., Gilar-Corbí, R., Izquierdo, A., and Castejón, J. L. (2020). Teacher training can make a difference: tools to overcome the impact of COVID-19 on primary schools. An experimental study. *Int. J. Environ. Res. Public Health* 17:8633. doi: 10.3390/ijerph17228633
- Pradhan, S., and Madihally, S. V. (2022). Teaching process simulators during COVID-19 pandemic: analysis on the digitalization of a dry laboratory. *J. Chem. Educ.* 99, 3007–3019. doi: 10.1021/acs.jchemed.2c00494
- R Core Team (2022) 'R: A language and environment for statistical computing.' Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Reese, H. W. (2011). The learning-by-doing principle. *Behav. Dev. Bull.* 17, 1–19. doi: 10.1037/h0100597
- Rosli, M. S., and Saleh, N. S. (2022). Technology enhanced learning acceptance among university students during Covid-19: integrating the full spectrum of self-determination theory and self-efficacy into the technology acceptance model. *Curr. Psychol.*, 1–20. doi: 10.1007/s12144-022-02996-1
- Sahu, I., and Majumdar, D. (2017) 'Detecting factual and non-factual content in news articles', in Proceedings of the Fourth ACM IKDD Conferences on Data Sciences. CODS '17: Fourth ACM IKDD Conferences on Data Sciences, Chennai India: ACM, pp. 1–12.
- Salas-Pilco, S. Z., Yang, Y., and Zhang, Z. (2022). Student engagement in online learning in Latin American higher education during the COVID-19 pandemic: a systematic review. *Br. J. Educ. Technol.* 53, 593–619. doi: 10.1111/bjet.13190
- Scherak, L., and Rieckmann, M. (2020). Developing ESD competences in higher education institutions—staff training at the University of Vechta. *Sustainability* 12:10336. doi: 10.3390/su122410336
- Seale, J. (ed.) (2020) *Improving accessible digital practices in higher education: Challenges and new practices for inclusion*. Cham: Springer International Publishing.
- Serianni, B. A., and Coy, K. (2014). Doing the math: supporting students with disabilities in online courses. *Teach. Except. Child.* 46, 102–109. doi: 10.1177/0040059914528330
- Shen, C., and Ho, J. (2020). Technology-enhanced learning in higher education: a bibliometric analysis with latent semantic approach. *Comput. Hum. Behav.* 104:106177. doi: 10.1016/j.chb.2019.106177
- Silge, J., and Robinson, D. (2016). Tidytext: text mining and analysis using tidy data principles in R. *J. Open Source Softw.* 1:37. doi: 10.21105/joss.00037
- Sonsona, D. A., Talidong, K. J. B., and Toquero, C. M. D. (2021). Game-based learning: reinforcing a paradigm transition on pedagogy amid COVID-19 to complement emergency online education. *Int. J. Didact. Stud.* 2, –10458. doi: 10.33902/IJODS.2021269730
- Tabatabai, S. (2020). Simulations and virtual learning supporting clinical education during the COVID 19 pandemic. *Adv. Med. Educ. Pract.* 11, 513–516. doi: 10.2147/AMEP.S257750
- Tan, L., and Kim, B. (2015). "Learning by doing in the digital media age" in *New media and learning in the 21st century*. eds. T.-B. Lin, V. Chen and C. S. Chai (Singapore: Springer Singapore (Education Innovation Series)), 181–197.
- Teixeira, A. F., Gonçalves, M. J. A., and Taylor, M. D. L. M. (2021). How higher education institutions are driving to digital transformation: a case study. *Educ. Sci.* 11:636. doi: 10.3390/educsci11100636
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., et al. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proc. Natl. Acad. Sci.* 117, 6476–6483. doi: 10.1073/pnas.1916903117
- Tobias, S., Fletcher, J. D., and Wind, A. P. (2014). "Game-based learning" in *Handbook of research on educational communications and technology*. eds. J. M. Spectoret al. (Springer New York: New York, NY), 485–503.
- Utomo, M. N. Y., Sudayanto, M., and Saddhono, K. (2020). Tools and strategy for distance learning to respond COVID-19 pandemic in Indonesia. *Ingénierie des Systèmes d'Inform.* 25, 383–390. doi: 10.18280/isi.250314
- Vial, G. (2019). Understanding digital transformation: a review and a research agenda. *J. Strateg. Inf. Syst.* 28, 118–144. doi: 10.1016/j.jsis.2019.01.003
- Vikas, S., and Mathur, A. (2022). An empirical study of student perception towards pedagogy, teaching style and effectiveness of online classes. *Educ. Inf. Technol.* 27, 589–610. doi: 10.1007/s10639-021-10793-9
- Wang, Q., and Zhao, G. (2023). Exploring the influence of technostress creators on in-service teachers' attitudes toward ICT and ICT adoption intentions. *Br. J. Educ. Technol.*, 1–19. doi: 10.1111/bjet.13315
- Wati, I. F. Yuniawatika (2020) 'Digital game-based learning as a solution to fun learning challenges during the Covid-19 pandemic': in Proceedings of the 1st International Conference on Information Technology and Education (ICITE 2020). 1st International Conference on Information Technology and Education (ICITE 2020), Malang, Indonesia: Atlantis Press.
- Yang, Y., Liu, K., Li, M., and Li, S. (2022). Students' affective engagement, parental involvement, and teacher support in emergency remote teaching during the COVID-19 pandemic: evidence from a cross-sectional survey in China. *J. Res. Technol. Educ.* 54, S148–S164. doi: 10.1080/15391523.2021.1922104
- Yew, E. H. J., and Goh, K. (2016). Problem-based learning: an overview of its process and impact on learning. *Health Prof. Educ.* 2, 75–79. doi: 10.1016/j.hpe.2016.01.004
- Yu, J., Denham, A. R., and Searight, E. (2022). A systematic review of augmented reality game-based learning in STEM education. *Educ. Technol. Res. Dev.* 70, 1169–1194. doi: 10.1007/s11423-022-10122-y
- Zaoui, F., and Souissi, N. (2020). Roadmap for digital transformation: a literature review. *Procedia Comput. Sci.* 175, 621–628. doi: 10.1016/j.procs.2020.07.090
- Zhang, L., Carter, R. A. Jr., Qian, X., Yang, S., Rujimora, J., and Wen, S. (2022). Academia's responses to crisis: a bibliometric analysis of literature on online learning in higher education during COVID-19. *Br. J. Educ. Technol.* 53, 620–646. doi: 10.1111/bjet.13191
- Zhang, X., Xu, Y. Y., and Ma, L. (2023). Information technology investment and digital transformation: the roles of digital transformation strategy and top management. *Bus. Process. Manag. J.* 29, 528–549. doi: 10.1108/BPMJ-06-2022-0254
- Штыхно, Д. А., Константинова, Л. В., and Гагиев, Н. Н. (2020). Переход вузов в дистанционный режим в период пандемии: проблемы и возможные риски. *Открытое образование* 24, 72–81. doi: 10.21686/1818-4243-2020-5-72-81



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Ayesha Kanwal,
University of Glasgow, United Kingdom
Ijeoma John-Adubasim,
University of Plymouth, United Kingdom
Tareq Al Jaber,
University of Hull, United Kingdom

*CORRESPONDENCE

Juan Dempere
✉ jdempere@hct.ac.ae

RECEIVED 16 April 2023

ACCEPTED 23 August 2023

PUBLISHED 08 September 2023

CITATION

Dempere J, Modugu K, Hesham A and
Ramasamy LK (2023) The impact of ChatGPT
on higher education.

Front. Educ. 8:1206936.

doi: 10.3389/feduc.2023.1206936

COPYRIGHT

© 2023 Dempere, Modugu, Hesham and
Ramasamy. 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 impact of ChatGPT on higher education

Juan Dempere*, Kennedy Modugu, Allam Hesham and
Lakshmana Kumar Ramasamy

Higher Colleges of Technology, Ras Al Khaimah, United Arab Emirates

Introduction: This study explores the effects of Artificial Intelligence (AI) chatbots, with a particular focus on OpenAI's ChatGPT, on Higher Education Institutions (HEIs). With the rapid advancement of AI, understanding its implications in the educational sector becomes paramount.

Methods: Utilizing databases like PubMed, IEEE Xplore, and Google Scholar, we systematically searched for literature on AI chatbots' impact on HEIs. Our criteria prioritized peer-reviewed articles, prominent media outlets, and English publications, excluding tangential AI chatbot mentions. After selection, data extraction focused on authors, study design, and primary findings. The analysis combined descriptive and thematic approaches, emphasizing patterns and applications of AI chatbots in HEIs.

Results: The literature review revealed diverse perspectives on ChatGPT's potential in education. Notable benefits include research support, automated grading, and enhanced human-computer interaction. However, concerns such as online testing security, plagiarism, and broader societal and economic impacts like job displacement, the digital literacy gap, and AI-induced anxiety were identified. The study also underscored the transformative architecture of ChatGPT and its versatile applications in the educational sector. Furthermore, potential advantages like streamlined enrollment, improved student services, teaching enhancements, research aid, and increased student retention were highlighted. Conversely, risks such as privacy breaches, misuse, bias, misinformation, decreased human interaction, and accessibility issues were identified.

Discussion: While AI's global expansion is undeniable, there is a pressing need for balanced regulation in its application within HEIs. Faculty members are encouraged to utilize AI tools like ChatGPT proactively and ethically to mitigate risks, especially academic fraud. Despite the study's limitations, including an incomplete representation of AI's overall effect on education and the absence of concrete integration guidelines, it is evident that AI technologies like ChatGPT present both significant benefits and risks. The study advocates for a thoughtful and responsible integration of such technologies within HEIs.

KEYWORDS

ChatGPT, AI, chatbot, artificial intelligence, higher education, natural language processing, transformers, digital revolution

1. Introduction

On November 30, 2022, the AI based chatbot called ChatGPT (Chat Generative Pre-trained Transformer) was launched as a prototype by OpenAI and rapidly gathered media attention for its comprehensive and articulate responses to questions spanning many domains of technical and professional knowledge (GPT, 2022). ChatGPT is an AI-based natural language processing (NLP)

system proficient in mimicking human-like communication with the end user. This virtual assistant enables responding to inquiries and supporting activities like crafting emails, writing essays, generating software code, and so on (Ortiz, 2022). This AI-based tool was initially offered open to the public free of charge because the launched demo and research version GPT-3.5 was intended to allow widespread general experimentation to get reinforcement learning from human feedback to be incorporated in the next version of GPT-4 (Goldman, 2022).

ChatGPT is a conversational AI chatbot engineered by OpenAI, a collective of researchers and technologists focused on constructing AI securely and responsibly. OpenAI was founded in 2015 by a team of tech innovators, and it has received substantial funding from tech giants such as Microsoft, Amazon, and Alphabet. The development of ChatGPT builds upon the tremendous advancements in the field of NLP. The GPT architecture has seen several iterations, with each new version achieving superior language generation, accuracy, and speed performance. The chatbot has been acclaimed as a breakthrough in NLP and used in various contexts, including customer service, education, and healthcare. In the field of learning, ChatGPT has been employed as an educational aid, replying to pupils' questions, giving feedback, and helping virtual conversations. ChatGPT can also be a writing helper, aiding people create grammatically accurate and logical text.

ChatGPT is a product of the GPT architecture, a leading-edge NLP model conditioned on copious amounts of text information to generate language similar to humans (GPT, 2022). A transformer is a deep learning model proposed by Vaswani et al. (2017), which introduced a self-attention approach that allows for a differential weighting of each input data component.

The revolutionary approach of transformers has been considered the most recent breakthrough in AI. Indeed, Chance (2022) describes transformers as deep learning models that allow expressing inputs in natural language to generate outputs like translations, text summaries, grammar and writing style correction, etc. Bellapu (2021) highlights the singularity of transformers as the amalgamation of convolutional neural networks and recurrent neural networks, with advantages such as better accuracy, faster processing, working with any sequential data, and forecasting.

Since its 2022 launch, AI chatbots like ChatGPT have sparked concerns in education. While risks about students' independent thinking and language expression skills deteriorating exist, banning the tool from

academic institutions should not be the answer (Dwivedi et al., 2023). Teachers and professors are uneasy about potential academic fraud with AI-driven chatbots such as ChatGPT (Meckler and Verma, 2022). The proficiency of ChatGPT spans from assisting in scholarly investigations to finalizing literary compositions for learners (Roose, 2022; Shankland, 2022). However, students may exploit technologies like ChatGPT to shortcut essay completion, endangering the growth of essential competencies (Shrivastava, 2022). Coursera CEO Jeff Maggioncalda believes that ChatGPT's existence would swiftly change any education using written assessment (Alrawi, 2023).

To gauge the media impact since the launch of ChatGPT on Nov. 30, 2022, we compared Google user search interests using Google Trends. This web service displays the search volume of queries over time in charts across countries and languages – Figure 1 shows ChatGPT's overwhelming media impact since its November 30, 2022 launch. Interest in the AI-based app exceeded Ukraine's war, news concerning U.S. President Joe Biden, Bitcoin, and the S&P 500. The data depicted in the chart is in line with Libert (2023) findings, which show that the search interest for ChatGPT soared to 112,740%.

As a result of the significant global breakthrough represented by the launch of ChatGPT, thousands of tech leaders and researchers, including Elon Musk, have called for a pause in the development of AI systems more potent than GPT-4 for 6 months or more, during which a set of shared safety protocols should be developed and implemented. An open letter of +50 K signatories emphasizes the need for robust AI governance systems, such as new regulatory authorities, tracking systems, auditing and certification, and liability for AI-caused harm. Finally, they suggest that a pause on AI development is necessary to ensure it is used for the benefit of all and to give society a chance to adapt (Bengio et al., 2020). The call comes as tech companies race to develop and deploy more powerful AI tools in their products, leading to concerns about biased responses, misinformation, privacy, and the impact on professions and relationships with technology.

On December 5, 2022, Altman (2022), the head of OpenAI, announced via Twitter that ChatGPT had garnered over a million users in under a week since its launch. The remarkable success of the Silicon Valley-based OpenAI has allowed it to forecast \$200 million in revenue in 2023 and \$1 billion by 2024, which placed the company's valuation at \$20 billion in a secondary share sale by the end of 2022 (Dastin et al., 2022). A more evident appreciation of OpenAI's

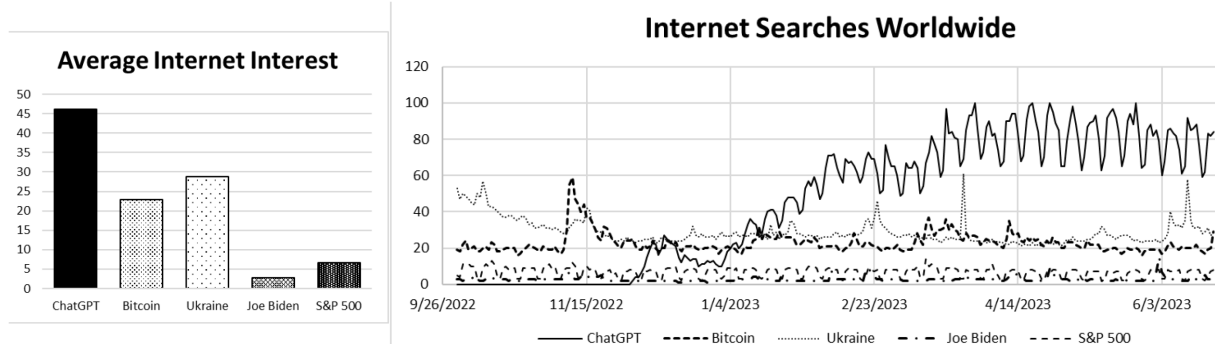


FIGURE 1

Search interest based on Google trends. The figures indicate search interest compared to the maximum point on the graph for the specified area and duration. A score of 100 signifies the zenith of popularity for the phrase. A rating of 50 implies the word is only half as prevalent. A value of 0 indicates insufficient data for the given term.

achievement can be gained by comparing it with those of other successful firms such as Netflix (+177 weeks), Facebook (+43 weeks), and Instagram (+10 weeks) in reaching the 1 million users mark (ColdFusion, 2022). Similarly, Hu (2023) reports that ChatGPT had achieved 100 million monthly active users within 2 months of its unveiling, making it the fastest-growing consumer application in recorded history, according to a UBS analysis.

ChatGPT4 was launched on March 14, 2023, and provides makers, developers, and creators with a powerful tool to generate labels, classify visible features, and analyze images. Compared to GPT-3.5, ChatGPT4 is more dependable, imaginative, and interactive and can tackle longer passages in one request because of the expanded setting length. Moreover, GPT-4 can handle textual and visual prompts and give both back, although the capacity to employ picture input is yet to be made available to the public. Furthermore, GPT-4 is more than 85% accurate in 25 languages, including Mandarin, Polish, and Swahili, and can write code in all major programming languages. Microsoft has brought out its Bing AI chatbot equipped with GPT-4 (Elecrow, 2023).

Recent news already provides information about the subsequent versions of ChatGPT. Indeed, Smith (2023) informs that OpenAI is working on the next major software upgrade for ChatGPT, GPT-5, which is expected to launch in winter 2023. If a report about the GPT-5 abilities is correct, it could bring ChatGPT to the point of AGI, making it indistinguishable from a human. OpenAI expects the intermediate ChatGPT version of GPT-4.5 to be launched in September or October 2023 if GPT-5 cannot be ready at that time (Chen, 2023).

Since its inception on November 30, 2022, ChatGPT has incited a notable amount of research articles. The deluge of scholarly works increases daily, making it unfeasible to offer an updated overview of the papers written regarding ChatGPT without becoming outdated in a few days or weeks. Some examples of such articles include Zhai (2023), who established that ChatGPT could resolve the most challenging issues in science education through automated assessment production, grading, guidance, and material suggestion. Similarly, Lund and Agbaji (2023) find that interest in using ChatGPT to benefit one's community was associated with information and privacy literacy but not data literacy among four northern Texas county residents.

Similarly, Sunjak (2022) findings suggest that ChatGPT can successfully replicate human-written text, raising doubt about the security of online tests in tertiary education. Likewise, Biswas (2023a) suggests that ChatGPT can be used to help improve the accuracy of climate projections through its ability to generate and analyze different climate scenarios based on a wide range of data inputs, including model parameterization, data analysis and interpretation, scenario generation, and model evaluation. Equally, Biswas (2023b) underlines the power of OpenAI's language model ChatGPT to advise people and groups in forming prudent judgments concerning their health and probes the potential applications of this chatbot in public health, as well as the upsides and downsides of its implementation. In the same way, Sobania et al. (2023) evaluated ChatGPT's proficiency at fixing bugs on QuixBugs and concluded that it was equivalent to CoCoNut and Codex - two widely used deep learning approaches - and was superior to typical program repair methods.

Likewise, Pavlik (2023) illustrates the potential and boundaries of ChatGPT by co-creating a paper with it and provides musings on the effects of generative AI on journalism and media education. Jeblick et al. (2022) conducted a probing analysis of 15 radiologists who asked about the quality of radiology reports produced by ChatGPT. Most

radiologists agreed that the simplified reports were precise, thorough, and risk-free. Still, a few misstatements, missed medical particulars, and potentially detrimental segments were noticed. Equally, Gao et al. (2022) tested ChatGPT by generating research abstracts from titles and journals in 10 high-impact medical journals ($n=50$). AI output detector identified the most generated abstracts (median of 99.98%) with a 0.02% probability of AI-generated output in original abstracts. Human reviewers identified 68% of generated abstracts but mistook 14% of original abstracts for generated.

Additionally, Chavez et al. (2023) suggest a neural network approach to forecast student outcomes without relying on personal data like course attempts, average evaluations, pass rates, or virtual resource utilization. Their method attains 93.81% accuracy, 94.15% precision, 95.13% recall, and 94.64% F1-score, enhancing the educational quality and reducing dropout and underperformance. Likewise, Kasepalu et al. (2022) find that an AI assistant can help teachers raise awareness and provide a data bank of coregulation interventions, likely leading to improved collaboration and self-regulation.

Patel and Lam (2023) discuss the potential use of ChatGPT, an AI-powered chatbot, for generating discharge summaries in healthcare. They report that ChatGPT allows doctors to input specific information and develop a formal discharge summary in seconds. Qin et al. (2023) analyze the ability of ChatGPT to perform zero-shot learning on 20 commonly used NLP datasets across seven categories of tasks. The researchers discovered that while ChatGPT excels in jobs requiring reasoning skills, it encounters difficulties performing specific tasks such as sequence tagging.

Generative Pre-trained Transformers have been used for research purposes in many areas, including climate (Alerskans et al., 2022), stock market (Ramos-Pérez et al., 2021), traffic flow (Reza et al., 2022), and flooding (Castangia et al., 2023). Additional examples of transformers being used for research purposes include predictions of electrical load (L'Heureux et al., 2022), sales (Vallés-Pérez et al., 2022), influenza prevalence (Wu et al., 2020), etcetera. Specifically, Lopez-Lira and Tang (2023) discovered that ChatGPT could accurately forecast stock market returns and surpasses traditional sentiment analysis approaches. They recommend integrating advanced language models into investment decision-making to enhance the accuracy of predictions and optimize quantitative trading strategies.

The fundamental purpose of this study is to deliver a qualitative analysis of the impact of AI chatbots like ChatGPT on HEIs by performing a scoping review of the existing literature. This paper examines whether AI chatbots can be used to enhance learning experiences and their potentially detrimental effect on the education process. Furthermore, this paper explores potential solutions to the prospective issues related to AI chatbots adopted by HEIs. Ultimately, this paper examines the existing literature on the current state of AI chatbot technology and its potential implications for future academic usage.

The novel contribution of this study resides in its comprehensive analysis of the impact of AI chatbots, particularly ChatGPT, on HEIs, synthesized through a detailed scoping review of existing literature. The primary research questions that drive this investigation include:

1. In what ways might AI chatbots like ChatGPT potentially replace humans in academic tasks, and what are the inherent limitations of such replacement?
2. How might AI technology be harnessed to detect and deter academic fraud?

3. What are the potential risks associated with the implementation of AI chatbots in Higher Education Institutions (HEIs)?
4. What academic activities in HEIs could be potentially enhanced with the adoption of AI chatbots like ChatGPT?
5. How might AI chatbots impact the digital literacy of students and their anxiety regarding AI technology?
6. What societal and economic implications might result from the wide-scale adoption of AI chatbots?

These questions guide the study's objectives, which include conducting a comprehensive review of existing literature to understand the current state of research, identifying trends and gaps in the literature, and informing future directions in the study of AI chatbots in HEIs. Additionally, the article highlights some critical societal and economic implications of AI adoption in HEIs, explores potential approaches to address the challenges and harness the benefits of AI integration, and underscores the need for strategic planning and proactive engagement from educators in leveraging AI technologies. This study uniquely amalgamates varied perspectives on the impact of AI chatbots in higher education, offering a broad, balanced, and nuanced understanding of this complex issue. In doing so, it aims to contribute significantly to the existing knowledge of AI in education and guide future research and policy-making in this rapidly evolving field.

2. Methods

2.1. Search strategy and databases

We used several databases to comprehensively cover the body of literature related to the impact of AI chatbots on higher education institutions. They include PubMed, Web of Science, IEEE Xplore, Scopus, Google Scholar, ACM Digital Library, ScienceDirect, JSTOR, ProQuest, SpringerLink, EBSCOhost, and ERIC. These databases were chosen due to their extensive coverage of scientific and scholarly publications across various disciplines, including technology, computer science, artificial intelligence, and education. Our search string was designed based on recent literature reviews of AI chatbots in HEIs (Okonkwo and Ade-Ibijola, 2021; Rahim et al., 2022). Our search strategy was systematic, combining relevant keywords and Boolean operators. Keywords included "ChatGPT," "AI chatbot," "Artificial Intelligence," "chatbot in education," "impact of AI chatbots on higher education," and their variants. Our search strategy was refined to ensure that it yielded the most relevant articles for our scoping review (Peters et al., 2015).

2.2. Inclusion and exclusion criteria

To streamline the process and maintain the quality and relevance of the study, we set out explicit inclusion and exclusion criteria. Our inclusion criteria included: (I) Published peer-reviewed articles that focus on the impact of AI chatbots, precisely like ChatGPT, on higher education institutions (HEIs). (II) Articles published in top media news outlets like the Washington Post, Forbes, The Economist, The Wall Street Journal, etc. (III) Studies that provide qualitative and quantitative evidence on using AI chatbots in HEIs. (IV) Articles published in English. (V) Conference proceedings and book chapters. Our exclusion criteria included articles that only tangentially mention AI chatbots or HEIs, without focusing on

the intersection of the two. In addition, secondary sources not published in English were also excluded.

2.3. Data extraction and analysis

Data extraction was performed once the final selection of articles was decided based on the inclusion and exclusion criteria. We extracted critical information from each document, such as authors, publication year, study design, the specific chatbot in focus, the context of use in HEIs, primary findings, and conclusions. Data analysis was guided by a narrative synthesis approach due to the various studies involved (O'Donovan et al., 2019). We analyzed the data both descriptively and thematically. The descriptive research focused on the bibliometric characteristics of the studies, including the number of studies, countries of origin, publication years, and the specific AI chatbots under investigation (Peters et al., 2020). The thematic analysis involved categorizing the findings into themes based on familiar patterns, such as specific applications of AI chatbots in HEIs, their benefits, limitations, ethical concerns, and future research directions. This systematic approach ensured that our scoping review was rigorous and adequately captured the state of research on the impact of AI chatbots on higher education institutions.

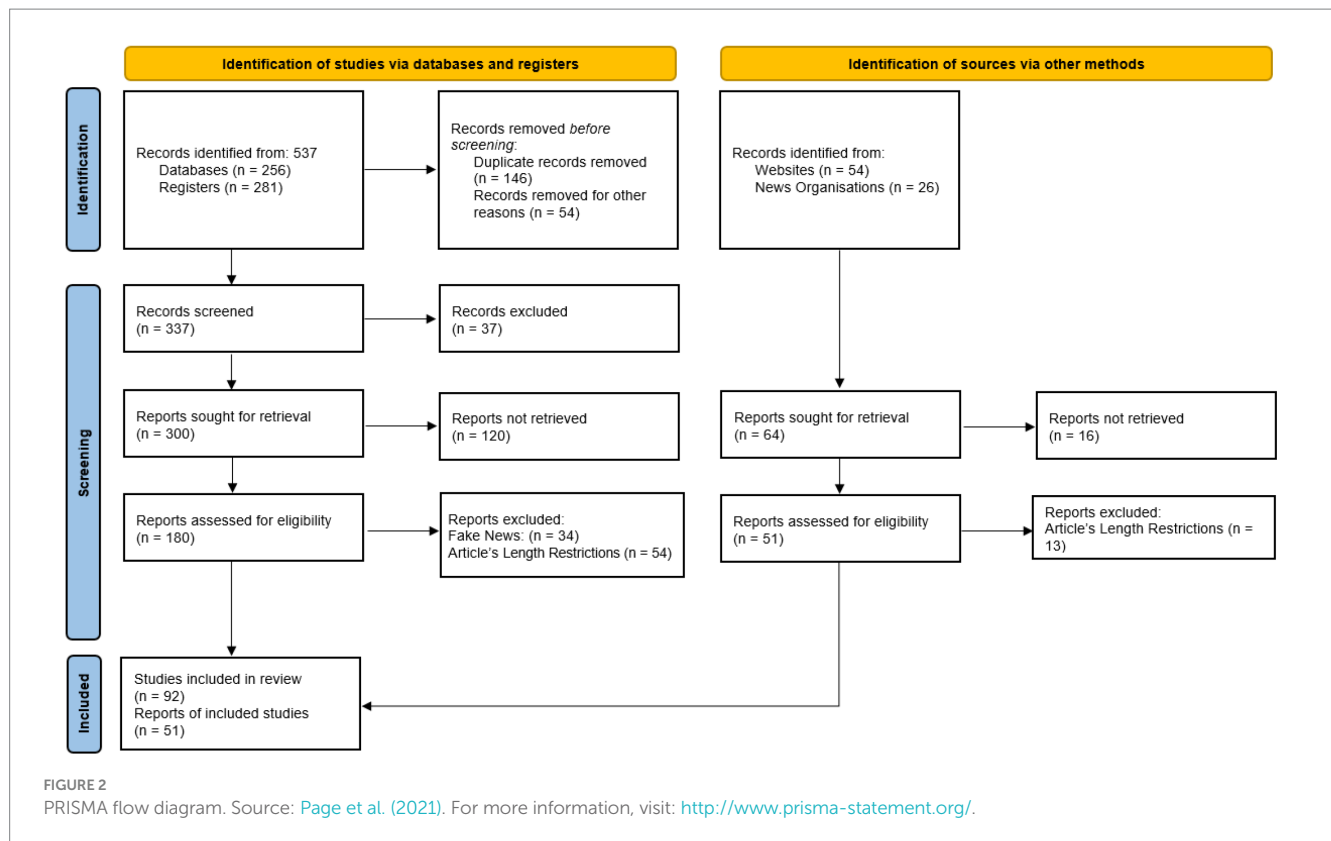
3. Results

Figure 2 shows the articles initially identified, those excluded based on title and abstract, and those excluded based on full-text review. It also shows the number of papers included in the final analysis and the reasons for exclusion at each stage. In the first segment, "Identification of studies via other methods," 80 records were identified, including 54 from various websites and 26 from organizations. Of these, 64 papers were sought for retrieval, while the remaining 16 were not retrieved for not satisfying the inclusion criteria. All 51 retrieved records were assessed for eligibility, and 13 were excluded due to the articles' length restrictions, leaving 10 to be included in the review.

In the second segment, "Identification of studies via databases and registers," 537 records were initially identified through various databases (256 documents) and registers (281 records). Before the screening, 200 papers were removed: 146 for being duplicates and 54 for other reasons. This screening left 337 records to be screened, of which 37 were excluded for not satisfying the inclusion criteria. Following the initial screening, 300 papers were sought for retrieval, but 120 were not retrieved according to the exclusion criteria. The remaining 180 records were assessed for eligibility, out of which 88 were excluded: 34 for being identified as fake news and 54 due to the article's length restrictions. This selection resulted in a final 92 articles being included in the review from the databases and registers, in addition to the 54 from other methods, which comprised 143 secondary sources for our analysis.

3.1. Potential approaches and inherent limitations of AI chatbots deployed to replace humans in HEIs

The concern about AI chatbots like ChatGPT replacing human beings to carry out a wide variety of tasks was expressed by The



Washington Post - Editorial Board (2022), who warn that the future of AI technology significantly more potent than today's will result in a price decline of many kinds of labors up to zero. Nevertheless, AI cannot undertake several academic tasks, including creative activities, such as inventing new courses or developing inventive teaching methods, and interpersonal interaction, such as counseling, providing personalized feedback, and resolving student issues. Additional tasks challenging to be substituted by AI comprise complex reasoning and problem-solving like selecting research projects or evaluating the effects of policy decisions, and empathy and understanding, such as coaching and providing emotional assistance (OpenAI, 2022).

Several articles support this view, including Murtarellia et al. (2021), who argue that chatbots lack valuable human traits like empathy, judgment, and discretion. Likewise, Felix (2020) warns that AI should not replace teachers since they can bring to the classroom a unique contribution that no machine can provide: their humanity. He argues that no AI application can provide valuable learning experiences regarding ethical norms and values, existential reflection, or a sense of self, history, and society. Equally, Brito et al. (2019) assert that some scholars believe AI will not supplant professors; however, they warn about the unavoidable reality of an existing AI-based technology that allows teaching-learning interactions without human intervention. This possibility represents a desirable attractive low-cost alternative, particularly for the private players in the education sector.

3.2. Harnessing AI technology to detect and deter academic fraud

The most effective technologies to tackle the challenges posed by AI chatbots like ChatGPT include AI-based plagiarism detection, text

similarity detection, and deep learning-based plagiarism detection, as well as online testing platforms such as ProctorU and ExamSoft for remote exams and academic fraud detection. Other technologies include digital examination, predictive analytics, machine learning for cheating detection, blockchain for secure student data, biometric verification for authentication, and digital rights management for IP protection (GPT-2 Output Detector Demo, 2022).

The development of AI-based plagiarism detection tools is supported by factual evidence. Indeed, since the launch of ChatGPT, the list of Internet resources for AI-generated content detection tools and services has been growing weekly (OpenAI, 2023; Originality, 2023; Allen Institute for AI, n.d.; Crossplag, n.d.; Writer, n.d.).

Regarding the online testing platforms mentioned above, Hu (2020) identifies several AI-based applications for this purpose, like ProctorU, Proctorio, and ProctorTrack. He argues that these online testing platforms analyze video recordings to determine suspicious student behavior, including interruptions of people entering the test room and the test taker's head or eye movements. Walters (2021) outlines issues associated with the widespread utilization of online proctoring software in New Zealand universities. He emphasizes the hardships students from disadvantaged backgrounds face, who may reside with extended family or in crowded housing. He may be disproportionately flagged due to unavoidable ambient noise, conversations, or people entering their exam room. In addition, those with disabilities or neurodiversity may be disadvantaged by AI-based surveillance of their movements and gaze.

The use of predictive analytics to detect academic fraud is also supported by academic research. Indeed, Trezise et al. (2019) confirm that keystroke and clickstream data can distinguish between authentically written pieces and plagiarized essays. Similarly, Norris (2019) explores strategies to thwart academic web fraud, such as

predictive analytics systems. He analyzes students' data from their interactions within their learning environments, including device details, access behavior, locations, academic advancement, *etcetera*, attempting to foretell students' behavioral trends and habits to detect questionable or suspicious activities.

Several academic articles also support using ML algorithms to detect cheating by analyzing student data. Some examples include Kamalov et al. (2021), who propose an ML approach to detect instances of student cheating based on recurrent neural networks combined with anomaly detection algorithms and find remarkable accuracy in identifying cases of student cheating. Similarly, Ruipérez-Valiente et al. (2017) employed an ML approach to detect academic fraud by devising an algorithm to tag copied answers from multiple online sources. Their results indicated high detection rates (sensitivity and specificity measures of 0.966 and 0.996, respectively). Equally, Sangalli et al. (2020) achieve a 95% generalization accuracy in classifying instances of academic fraud using a Support Vector Machine algorithm.

The use of blockchain for data tampering prevention is also supported by academic research. Reis-Marques et al. (2021) analyzed 61 articles on blockchain in HEIs, including several addressing educational fraud prevention. Tsai and Wu (2022) propose a blockchain-based grading system that records results and activities, preventing post-grade fraud. Islam et al. (2018) suggest a two-phase timestamp encryption technique for question sharing on a blockchain, reducing the risk of exam paper leaks and maintaining assessment integrity.

The use of biometric verification for cheating prevention is also backed by research. Rodchua et al. (2011) review biometric systems, like fingerprint and facial recognition, to ensure assessment integrity in HEIs. Similarly, Agulla et al. (2008) address the lack of face recognition in learning management systems and propose a FaceTracking application using webcam video. Agarwal et al. (2022) recommend an ML-based keystroke biometric system for detecting academic dishonesty, reporting 98.4% accuracy and a 1.6% false-positive rate.

3.3. Potential risks of implementing AI Chatbots in HEIs

Adopting AI chatbots in HEIs presents various risks, such as privacy breaches, unlawful use, stereotyping, false information, unexpected results, cognitive bias, reduced human interaction, limited accessibility, and unethical data gathering (OpenAI, 2022). Indeed, Baidoo-Anu and Ansah (2023) emphasize certain inherent drawbacks of the chatbot, such as misinformation, augmenting preexisting prejudices through data training, and privacy concerns. Similarly, Akgun and Greenhow (2022) caution against using AI-based algorithms to predict individual actions from chatbot-human interaction information gathered, raising questions regarding fairness and self-freedom. Likewise, Murtarellia et al. (2021) draw attention to the increased information asymmetry from AI chatbots such as ChatGPT, indicating that human conversations with these bots can enable the collection of personal data to build a user profile. AI chatbots can identify patterns that create an informational advantage for the algorithm's owner. For example, some HEIs may leverage chatbots to sway students' attitudes toward academic advice to artificially boost enrollment in specific courses to the detriment of others.

In the same way, Miller et al. (2018) cautioned about the potential perils of using social data, including human prejudice to train AI systems, which could lead to prejudicial decision-making processes. Similarly, Akgun and Greenhow (2022) inform the risks of adopting AI-based technologies in academia, including the likely preservation of prevailing systemic bias and discrimination, the perpetuation of unfairness for students from historically deprived and marginalized groups, and magnification of racism, sexism, xenophobia, and other practices of prejudice and injustice. They also advise about the AI-based systems capable of monitoring and tracking students' thoughts and ideas, which may result in surveillance systems capable of threatening students' privacy.

Regarding the negative impact of replacing human interaction in the learning process in terms of engagement and learning outcomes, Fryer et al. (2017) study the chatbots' long-term effects on task and course interest among foreign language students and find that a significant decline in students' task interest when interacting with a chatbot but not a human partner. Regarding the risk of misinformation, Bushwick and Mukerjee (2022) suggest that AI chatbots should be subject to some form of regulation due to the risks associated with a technology capable of human-like writing and answering to a wide range of topics with advanced levels of fluency and coherence. These risks include spreading misinformation or impersonating individuals. Regarding the unanticipated outcomes of AI chatbots, they are referred to as hallucinations: unpredictable AI outputs caused by data beyond its training set. Additional information and references about this issue are provided later in this article.

The issue of stereotyping has mixed academic evidence. For example, Bastiansen et al. (2022) deployed the Stereotype Content Model to research the effects of warmth and gender of a chatbot on stereotypes, trustworthiness, aid, and capability. They find no divergent outcomes stemming from exposure to heat and assigned gender. Alternatively, Leavy (2018) argues that machine intelligence reflects gender biases in its data. Although attempts have been made to address algorithmic bias, they still need to pay more attention to the role of gender-based language. Women, who are leading this field, are best positioned to identify and solve this issue. Achieving gender parity in ML is crucial to prevent algorithms from perpetuating harmful gender biases against women.

Finally, the issue of accessibility has barely been analyzed in the academic literature. Stanley et al. (2022) identify 17 distinct sources resulting in 157 different suggestions for making a chatbot experience that is accessible, which they grouped into five groups: content, user interface, integration with other web content, developer process & training, and testing.

3.4. Academics activities that may be enhanced by the adoption of AI chatbots in HEIs

Adopting AI chatbots like ChatGPT in HEIs can positively affect various academic activities, including admissions, as they can streamline enrollment with tailored approaches to individual student needs. Student services can also benefit from AI chatbots, as they can provide personalized assistance with financing, scheduling, and guidance. Additionally, AI chatbots can enhance teaching by creating interactive learning experiences to assist students in comprehending

course material, providing personal feedback, and aiding researchers in data collection and analysis. Furthermore, AI chatbots can improve student life by furnishing students with personalized support for events and activities, advice on student life, and social interaction. Lastly, AI chatbots can increase student retention by providing customized advice and assistance (OpenAI, 2022).

Regarding the benefits of AI on admissions, Page and Gehlbach (2017) assess the efficiency of a conversational AI system to assist first-year students transitioning to college through personalized text message-based outreach at Georgia State University. Their findings reveal improved success with pre-enrollment requirements and timely enrollment among study participants. Arun et al. (2019) also assess an AI-based CollegeBot's effectiveness in providing students with university-related information, class schedules, and assessment timetables. Their study substantiates the advantages of chatbots for student services. Likewise, Slepankova (2021) finds that AI chatbot applications enjoying significant student support include delivering course material recap, study material suggestions, and assessment requirements information.

Georgescu (2018) and other academic articles suggest that chatbots can transform education by supporting content delivery and assessment on various topics, including multimedia content and AI-based speeches. Similarly, Essel et al. (2022) studied the adoption's impact of a virtual teaching assistant in Ghanaian HEIs, finding students who interacted with the chatbot had higher academic performance than those who interacted with the course instructor. Wang et al. (2017) investigated the impact of chatbots in immersive virtual English learning environments, discovering this tech enhances students' perception of such settings. Kerly and Bull (2006) studied chatbots' benefits in developing university students' negotiation skills. Tegos et al. (2015) analyzed the effects of chatbots in collaborative learning experiences among college students, finding that tech increases various knowledge acquisition measures. Lastly, Shorey et al. (2019) examined the benefits of using a chatbot as a virtual patient to develop nursing students' communication skills, finding this technology improves students' perceived self-efficacy and trust in their abilities.

The present article constitutes an excellent first example regarding the benefits of chatbots in research. However, additional examples would include studies analyzing the influence of AI chatbots among university students experiencing symptoms of depression and anxiety (Fitzpatrick et al., 2017; Fulmer et al., 2018; Klos et al., 2021). Similarly, Bendig et al. (2019) develop a comprehensive literature review on using chatbots in clinical psychology and psychotherapy research, including studies employing chatbots to foster mental health. Likewise, Dwivedi et al. (2023) discuss the impact of ChatGPT on academic research, noting its potential to improve the quality of writing and make research more accessible to non-experts while also posing challenges such as the authenticity and reliability of generated text and accountability and authorship issues.

Additionally, several articles report using AI chatbots for gathering qualitative information for research purposes. Some examples include Xiao et al. (2020), who create a prototype to generate two chatbots – one with active listening skills and one without – and evaluate both chatbots using 206 participants to compare their performance and conclude that their study provides practical methods for building interview chatbots effectively. Similarly, Nunamaker et al. (2011) suggest gathering human physiology and behavior information

during interactions with chatbot-like technology. Pickard et al. (2017) compare the qualitative data collected from automated virtual interviewers, called embodied conversational agents, versus the information obtained by human interviewers. Tallyn et al. (2018) use a chatbot to gather ethnographic data for analysis. Xiao et al. (2020) assess the effectiveness of the limitations of chatbots in conducting surveys. Finally, Kim et al. (2019) find that chatbot-based surveys can produce higher-quality data than web-based surveys.

Concerning the use of AI chatbots to retain students, earlier articles highlight the advantages these chatbots offer, potentially improving student retention. Indeed, Lee et al. (2022) investigate a computer-generated conversational agent-aided evaluation system and realize that it advances student-achievement results, including scholarly accomplishment, assurance, learning mentality, and enthusiasm. They infer that chatbots can heighten learner participation in the educational process. Other articles analyze the benefits of chatbots to provide students with standardized academic information, like course content (Cunningham-Nelson et al., 2019), practice exercises and questions (Sinha et al., 2020), frequently asked college questions (Ranoliya et al., 2017; Clarizia et al., 2018), assessment criteria (Benotti et al., 2018; Durall and Kapros, 2020), assignment calendars (Ismail and Ade-Ibijola, 2019), etcetera.

Previous research works encompass the analysis of chatbots dedicated to informing about campus physical locations (Mabunda and Ade-Ibijola, 2019), teaching computer programming concepts (Pham et al., 2018; Zhao et al., 2020), providing academic and administrative services (Hien et al., 2018), etcetera. Equally, Sandu and Gide (2022) study the benefits of chatbots in the Indian educational sector and find that this technology can improve communication, learning, productivity, and teaching assistance effectiveness and minimize interaction ambiguity. Lastly, Aldhaen (2022) suggests that implementing AI in the academic world will improve educational and non-academic operations governance.

3.5. Impact of AI chatbots on the digital literacy of students and their AI-caused anxiety

The implications of the launch of ChatGPT refer primarily to proactive approaches to face the academic integrity challenges posed by AI chatbots like ChatGPT. Meckler and Verma (2022) suggest requiring students to write by hand during class sessions to ensure successful monitoring efforts. Alternatively, Shrivastava (2022) emphasizes the relevance of teaching digital literacy early on to allow students critically assess the source of the information they receive.

Digital literacy should also teach students the risks of relying on AI-based technologies. These risks include hallucinations: AI-generated responses not explained by training data. Several authors have studied AI-generated hallucinations. Indeed, Cao et al. (2017) find that 30% of the outputs generated by state-of-the-art neural summarization applications suffer from hallucination problems. Similarly, Falke et al. (2019) study the most recent technologically advanced summarization systems and find that they produce about 25% of hallucination errors in their summaries. Likewise, Maynez et al. (2020) see 70%+ of single-sentence summaries show intrinsic/extrinsic hallucinations in AI-based systems (Recurrent, Convolutional, and Transformers).

Shuster et al. (2021) also studied neural retrieval in loop architectures. They found they enabled open-domain conversational capabilities, including generalizing scenarios not seen in training and reducing knowledge hallucination in advanced chatbots. Equally, Bang et al. (2023) find that ChatGPT has 63.41% accuracy on average in 10 different reasoning categories under logical reasoning, non-textual reasoning, and commonsense reasoning, which makes it an unreliable reasoner. The authors also report that ChatGPT suffers from hallucination problems.

Finally, digital literacy training must cover the risk of plagiarism when using AI chatbots. Ghosal (2023) notes ChatGPT's downside of lacking plagiarism verification as it picks sentences from training data. King and chatGPT (2023) discuss AI and chatbots' history and potential misuse, particularly in higher education, where plagiarism is a growing concern. Professors can minimize cheating via ChatGPT using various assessment methods and plagiarism detection software (GPT-2 Output Detector Demo, 2022; OpenAI, 2023; Originality, 2023; Allen Institute for AI, n.d.; Crossplag, n.d.; Writer, n.d.).

Several public and private organizations have been alarmed by the launch of ChatGPT. Lukpat (2023) reveals that New York City schools blocked access to ChatGPT on its networks and devices due to fears that students could use the AI app to answer questions, do schoolwork, or write essays. Soper (2023) details that Seattle Public Schools is also prohibiting ChatGPT. Cassidy (2023) reports that Australian universities have had to adjust their approach to testing and grading due to fears of students using AI to write essays. They have set new rules stating that using AI is considered cheating. McCallum (2023) reports that Italy initially banned OpenAI's ChatGPT due to privacy issues by arguing that there is no legal reason to gather and store private data for training algorithms. OpenAI's lack of transparency about its architecture, model, hardware, computing, training, and dataset construction has caused further concern (Brodikin, 2023). Ryan-Mosley (2023) informs the European Parliament's endorsement of the preliminary guidelines of the EU AI Act, which stipulate barring the use of AI emotion-detection in specified areas, a possible banning real-time biometrics and predictive policing in public spaces, outlawing public agencies' social scoring, prohibiting copyrighted content in LLMs' training datasets, *etcetera*.

Gaceta (2023) reports Paris' Institute of Political Science banned students from using ChatGPT to prevent academic fraud. Academic journals updated policies to prohibit ChatGPT as an author (Thorp, 2023). Dwivedi et al. (2023) recommend forbidding ChatGPT or equivalent software from producing intellectual outputs. However, in the non-academic world, Bensinger (2023) reports over 200 ChatGPT-authored e-books on Amazon. Libert (2023) reveals that Study.com questioned both teachers and learners, discovering that 72% of instructors were worried about the repercussions of ChatGPT on plagiarism, yet only 34% felt it should be banned. They uncovered that 89% of pupils employed ChatGPT for homework aid, 48% for an assessment/quiz, 53% for composition, and 22% for a paper structure. The nature of ChatGPT and the need for author disclosure make determining the number of AI-generated e-books challenging.

AI chatbots pose security concerns, with potential risks including disinformation and cyberattacks. OpenAI's CEO, Sam Altman, is aware of the dangers but optimistic about the technology's benefits (Ordóñez et al., 2023). Check Point (2023) reports underground hacking communities using OpenAI to design malicious tools, and skilled threat actors will likely follow. Perry et al. (2022) conducted a

large-scale study on using an AI code assistant for security tasks and found that participants with AI access produced less reliable code.

ChatGPT is a powerful tool for revolutionizing the academic world, and fear of overthrowing the existing order has traditionally resulted in repressive, oppressive, and other drab strategies utilized by those who dread forfeiting the positions that the current system grants them. As has been the case with numerous other scientific and technological advances that have been banned throughout human history (including Darwin's theory of evolution, Copernicus's heliocentric model, specific immunizations, blood transfusions, etc.), likely, ChatGPT and similar AI-powered applications may soon suffer a similar fate.

3.6. Societal and economic implications from the wide-scale adoption of AI chatbots

Individuals are apprehensive of AI owing to its capacity to disrupt many industries and result in job loss. Furthermore, many are concerned that AI could become so advanced that it would take over human control and make decisions for us. The notion of machines and robots replacing humans in the workplace can be disconcerting. Moreover, some fear that a powerful AI could become so potent that it would endanger humanity (OpenAI, 2022).

The reasons for humans to fear the development of AI chatbots like ChatGPT are many and compelling, although it is too early to support such fears with solid statistical evidence. Therefore, when writing this article, only partial and anecdotal evidence can be presented. Indeed, according to a report by researchers at Stanford University (AI Index Steering Committee, 2023), 36% of experts believe that decisions made by AI could lead to "nuclear-level catastrophes" (AI Index Steering Committee, 2023, p. 337). While the majority of researchers surveyed believe AI could lead to a "revolutionary change in society" (AI Index Steering Committee, 2023, p. 337), they also warned of the potential dangers posed by technology development.

Some of the human fears of AI are derived from its capacity to replicate academic achievements that would require years of investment in time, money, and effort, in just a few seconds. Indeed, ChatGPT3.5 and ChatGPT4 both excelled on standardized exams such as the Uniform Bar Exam, GREs, SATs, USABO Semifinal Exam 2020, Leetcode coding challenges, and AP exams. ChatGPT4 outperformed ChatGPT3.5 in all difficulty levels of the Leetcode coding challenges and subjects like Biology, Calculus BC, Chemistry, Art History, English, Macroeconomics, Microeconomics, Physics, *etcetera* in AP exams. Additionally, ChatGPT4 performed better than ChatGPT3.5 on Medical Knowledge Self-Assessment Program Exam, USABO Semifinal Exam 202, USNCO Local Section Exam 2022, Sommelier exams, *etcetera* (OpenAI, 2023). Likewise, the OpenAI (OpenAI, 2023) GPT-4 Technical Report reveals that GPT-4 demonstrates proficiency comparable to humans on multiple tests, such as a fabricated bar exam that it obtained a rank in the uppermost tenth percentile of test participants. GPT-4 outperforms existing large language models (LLMs) on a collection of NLP tasks and exceeds most reported state-of-the-art systems.

AI replacing human jobs is a mortal fear. Goldman Sachs (Hatzius et al., 2023) predicts that ChatGPT and other generative AI could

eliminate 300 million jobs worldwide. Researchers estimate that AI could replace 7% of US employment, complement 63%, and leave 30% unaffected. AI's global adoption may boost GDP by 7% (Hatzius et al., 2023). Taulli (2019) suggests that automation technology will take over “repetitive processes” in fields like programming and debugging. Positions requiring emotional intelligence, empathy, problem-solving, critical decision-making, and adaptabilities, like social workers, medical professionals, and marketing strategists, are difficult for AI to replicate.

In the same way, Felten et al. (2023) assess the impact of AI language modeling advances on occupations, industries, and geographies and find that telemarketers and post-secondary teachers of English, foreign languages, and history are most exposed to language modeling. Additionally, legal services, securities, commodities, and investments are most exposed to language modeling advances. Similarly, Tate (2021) cautions that AI's rapid evolution could eradicate the “laptop class” of employees in the upcoming decade. At the same time, blue-collar vocations necessitating hands-on expertise and manual labor will remain safe. She further cautions of AI potentially supplanting white-collar jobs in law, finance, media, and healthcare. She advises those seeking job stability to pursue blue-collar roles instead of STEM fields and “knowledge economy” positions that will be obsolete in the imminent transformation.

Equally, Bubeck et al. (2023) argue that the emergence of GPT-4 and other large LLMs will challenge traditional notions of human expertise in various professional and scholarly fields. They suggest that the capabilities of GPT-4 may raise concerns about the potential for AI to displace or reduce the status of human workers in highly skilled professions. The rise of LLMs could also widen the “AI divide” between those with access to the most powerful AI systems and those without, potentially amplifying existing societal divides and inequalities.

Accordingly, GESTION (2023) reveals that, as established by a survey of 1,000 US business magnates by ResumeBuilder.com, virtually half of the businesses have already adopted ChatGPT, and a further 30% are looking to do so. It is supposed that 48% of ChatGPT customers have supplanted personnel and saved over \$75,000. Notwithstanding this, the dominant part of business magnates is content with the standard of ChatGPT's work, with 55% rating it ‘excellent’ and 34% rating it ‘very good.’ ChatGPT is employed for code writing (66%), content production (58%), customer service (57%), and document/meeting summaries (52%). It is also used to write job descriptions (77%), craft job interview applications (66%), and reply to job seekers (65%).

Similarly, Eloundou et al. (2023) evaluate the impact of GPTs and other LLMs on 19,262 tasks and 2,087 job processes from the O*NET 27.2 database. The study showed 80% of the US workforce may have 10% of their functions affected, and 19% may see 50% or more impacted. 15% of US worker tasks could be faster with LLMs, increasing to 47–56% with LLM-powered software. Science/critical thinking jobs are less affected, while programming/writing skills are more vulnerable. Higher-wage occupations have more LLM exposure, with no correlation to employment levels. Those with some college education but no degree have high LLM exposure. Findings suggest potential economic, social, and policy implications, requiring preparedness for disruption.

HEIs can use knowledge of AI's impact on the job market to adjust their curriculum, prioritizing skills AI cannot replicate, such as problem-solving and critical decision-making. Additionally, institutions can teach students to use and develop AI to their advantage, preparing them for the changing job market and ensuring their success in the workplace.

4. Discussion

4.1. Implications

The hope that the AI development race will pause, as suggested by tech leaders and researchers, including Elon Musk, for 6 months or more is naive. Oxford Insights' 2022 Government AI Readiness Index (Insights, 2022) ranks 160 countries on AI readiness for public services, with 30% having released a national AI strategy and 9% developing one. This finding highlights the importance of AI to leaders worldwide. Figure 3 shows the top 20 nations in the index.

Pausing in the AI development race will leave countries behind, and developed economies cannot afford to pay such a price. However, the emergence of ChatGPT and similar technologies may require regulatory frameworks to address privacy, security, and bias concerns, ensuring accountability and fairness in AI-based services. Rules must not impede AI-based tech development, as uncertainty can threaten investments. The US commerce department is creating accountability measures for AI tools (Bhuiyan, 2023), soliciting public feedback on assessing performance, safety, effectiveness, and bias, preventing misinformation, and ensuring privacy while fostering trustworthy AI systems.

Private HEIs will likely lead the AI revolution, driven by cost-saving, productivity, student satisfaction, and reputation. ChatGPT

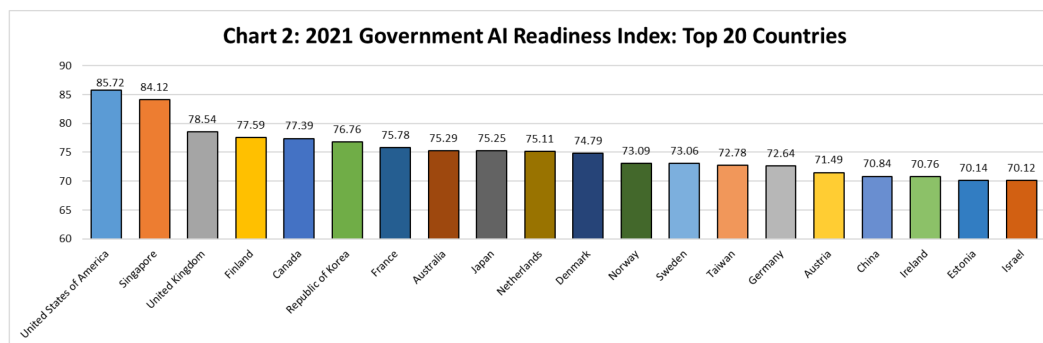


FIGURE 3

2021 government AI readiness index: top 20 countries.

can revolutionize education, enterprises, and linguistics, offering 24/7 access to virtual mentors with internationally recognized wisdom, fostering inventiveness, and providing discernment into consumer conduct (Dwivedi et al., 2023). Personalized learning experiences, immediate feedback, and language support are also possible with ChatGPT. Dwivedi et al. (2023) urge embracing digital transformation in academia and using ChatGPT to stimulate discussions about fundamental principles.

Faculty should proactively embrace AI chatbots such as ChatGPT as powerful teaching, research, and service tool. By becoming informed and trained on AI, they can learn its capabilities and limitations, identify assessment strategies to reduce academic fraud and create innovative pedagogical solutions for future developments. In an AI-driven world, traditional learning will soon become obsolete. Instead, students will query AI for answers to their problems, from cooking to coding. Even better, they could use AI tools to refine existing solutions and exercise their imagination to create new solutions to future challenges with endless possibilities.

AI-based learning experiences must recognize that AI technologies are trained using existing data and are ill-equipped to tackle novel problems without training data. For instance, AI might face challenges in dealing with the unprecedented obstacles humans may encounter during space exploration. Learning experiences must teach them to analyze cases with limited historical data to train available AIs. This example is one of the many approaches for adopting AI in the current academic world, which must shift rapidly to survive.

Another significant factor when adopting AI technologies into the existing learnings experiences at HEIs is exploiting the deficiencies of current and future AI-based technologies in terms of hallucinations, bias in training data leading to biased output, AI-generated ethical dilemmas, novelty security and privacy concerns, poor generalizability of AI models, lack of accounting for human context and understanding, *etcetera*. All these examples constitute opportunities to include AI-assisted curricula where the notion of AI-assisted student cheating becomes unappealing.

4.2. Limitations

Our study has several limitations. Firstly, the study is based on a scoping review of existing literature, which may not provide a complete or up-to-date picture of the effects of AI-based tools in the education sector. Moreover, the research relies on anecdotal evidence and partial data, limiting the findings' generalizability. Additionally, the study does not investigate the implementation challenges and practical implications of integrating AI chatbots into the HEIs' systems. Furthermore, the research does not consider the social and ethical implications of AI's increasing role in education, such as the impact on human connection and interpersonal skills development. Finally, the study does not provide concrete recommendations or guidelines for HEIs to integrate AI technologies into their teaching, research, and student services.

5. Conclusion

The primary contribution of this article is the development of qualitative research on the impact of AI chatbots like ChatGPT on

HEIs by employing a scoping review of the current literature. Developing AI-based tools such as ChatGPT increases the likelihood of replacing human-based teaching experiences with low-cost chatbot-based interactions. This possibility may result in biased teaching and learning experiences with reduced human connection and support. We also provide secondary source evidence that adopting AI-based technologies like ChatGPT can provide many benefits to HEIs, including increased effectiveness on student services, admissions, retention, *etcetera*, and significant enhancements to teaching and research activities. We also verify that the risks involved in adopting this technology in the education sector are substantial, including sensitive issues such as privacy and accessibility concerns, unethical use, data collection, misinformation, technology overreliance, cognitive bias, replacement of human interaction, etc.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

The authors of this academic research article declare that there were no funding sources for this research, nor any competing interests.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2023.1206936/full#supplementary-material>

References

- Agarwal, N., Danielsen, N., Graval, P., and Bours, P. (2022). Contract cheat detection using biometric keystroke dynamics. 20th international conference on emerging eLearning technologies and applications (ICETA). IEEE Xplore, 15–21
- Agulla, E., Rifón, L., Castro, J., and Mateo, C. (2008). Is my student on the other side? Applying biometric web authentication to E-learning environments. Eighth IEEE International Conference on Advanced Learning Technologies, IEEE Xplore, 551–553
- AI Index Steering Committee (2023). Artificial intelligence index report 2023. Stanford University, Stanford Institute for Human-Centered Artificial Intelligence (HAI). Retrieved from https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf
- Akgun, S., and Greenhow, C. (2022). Artificial intelligence in education: addressing ethical challenges in K-12 settings. *AI Ethics* 2, 431–440. doi: 10.1007/s43681-021-00096-7
- AlDhaen, F. (2022) in *The use of artificial intelligence in higher education – systematic review*. ed. M. Alaali (Cham: COVID-19 Challenges to University Information Technology Governance, Springer)
- Alerskans, E., Nyborg, J., Birk, M., and Kaas, E. (2022). A transformer neural network for predicting near-surface temperature. *Meteorol. App.* 29, 1–9. doi: 10.1002/met.2098
- Allen Institute for AI. (n.d.). Grover - a state-of-the-art defense against neural fake news. Available at: <https://grover.allenai.org/detect> (Accessed April 8, 2023).
- Alrawi, M. (2023). Davos 2023: AI chatbot to 'change education forever within six months'. The National News, January 17, 2023. Available at: <https://www.thenationalnews.com/business/2023/01/17/davos-2023-ai-chatbot-to-change-education-forever-within-six-months/> (Accessed January 22, 2023).
- Altman, S. (2022). ChatGPT launched on Wednesday. Today it crossed 1 million users! [tweet]. Twitter. Available at: <https://twitter.com/sama/status/1599668808285028353> (Accessed January 1, 2023).
- Arun, K., Nagesh, S., and Ganga, P. (2019). A multi-model and Ai-based Collegebot management system (Aicms) for professional engineering colleges. *Int. J. Innov. Technol. Explor. Eng.* 8, 2910–2914. doi: 10.35940/ijitee.18818.078919
- Baidoo-Anu, D., and Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. *SSRN Preprint (January 25, 2023)*. doi: 10.2139/ssrn.4337484
- Bang, Y., Cahyawijaya, S., Lee, N., Dai, W., Su, D., and Fung, P. (2023). A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity. *arXiv preprint arXiv:2302.04023*. doi: 10.48550/arXiv.2302.04023
- Bastiansen, M., Kroon, A., and Araujo, T. (2022). Female chatbots are helpful, male chatbots are competent? *Publizistik* 67, 601–623. doi: 10.1007/s11616-022-00762-8
- Bellapu, A. (2021). Why transformer model for natural language processing? Analytics Insight. Available at: <https://www.analyticsinsight.net/why-transformer-model-for-natural-language-processing/> (Accessed January 1, 2023).
- Bendig, E., Erb, B., Schulze-Thuesing, L., and Baumeister, H. (2019). The next generation: Chatbots in clinical psychology and psychotherapy to foster mental health—a scoping review. *Verhaltenstherapie* 29, 266–280. doi: 10.1159/000499492
- Bengio, Y., Russell, S., and Musk, E. (2020). Pause giant AI experiments: AN OPEN LETTER. Available at: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (Accessed April 1, 2023).
- Benotti, L., Martnez, M., and Schapachnik, F. (2018). A tool for introducing computer science with automatic formative assessment. *IEEE Trans. Learn. Technol.* 11, 179–192. doi: 10.1109/TLT.2017.2682084
- Bensinger, G. (2023). ChatGPT launches boom in AI-written e-books on Amazon. Reuter. Available at: <https://www.reuters.com/technology/chatgpt-launches-boom-ai-written-e-books-amazon-2023-02-21/> (Accessed March 4, 2023).
- Bhuiyan, J. (2023). 'We have to move fast': US looks to establish rules for artificial intelligence. The Guardian. Available at: <https://www.theguardian.com/technology/2023/apr/11/us-artificial-intelligence-rules-accountability-measures> (Accessed April 13, 2023).
- Biswas, S. S. (2023a). Potential use of chat GPT in global warming. *Ann. Biomed. Eng.* 51, 1126–1127. doi: 10.1007/s10439-023-03171-8
- Biswas, S. S. (2023b). Role of chat GPT in public health. *Ann. Biomed. Eng.* 51, 868–869. doi: 10.1007/s10439-023-03172-7
- Brito, C., Ciampi, M., Sluss, J., and Santos, H. (2019). Trends in engineering education: a disruptive view for not so far future. 2019 18th international conference on information technology based higher education and training (ITHET), 1–5
- Brodtkin, J. (2023). GPT-4 poses too many risks, and releases should be halted, AI group tells FTC. Available at: <https://arstechnica.com/tech-policy/2023/03/ftc-should-investigate-openai-and-halt-gpt-4-releases-ai-research-group-says/> (Accessed April 1, 2023).
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., and Zhang, Y. (2023). Sparks of artificial general intelligence: early experiments with GPT-4. *arXiv preprint arXiv:2303.12712*. doi: 10.48550/arXiv.2303.12712
- Bushwick, S., and Mukerjee, M. (2022). ChatGPT explains why AIs like ChatGPT should be regulated. Available at: <https://www.scientificamerican.com/article/chatgpt-explains-why-ais-like-chatgpt-should-be-regulated/> (Accessed January 1, 2023).
- Cao, Z., Wei, F., Li, W., and Li, S. (2017). Faithful to the original: fact aware neural abstractive summarization. *arXiv preprint arXiv:1711.04434v1 [cs.IR]*. doi: 10.48550/arXiv.1711.04434
- Cassidy, C. (2023). Australian universities to return to 'pen and paper' exams after students caught using AI to write essays. Available at: <https://www.theguardian.com/australia-news/2023/jan/10/universities-to-return-to-pen-and-paper-exams-after-students-caught-using-ai-to-write-essays> (Accessed January 18, 2023).
- Castangia, M., Medina, L., Aliberti, A., Rossi, C., Macii, A., Macii, E., et al. (2023). Transformer neural networks for interpretable flood forecasting. *Environ. Model. Softw.* 160:105581. doi: 10.1016/j.envsoft.2022.105581
- Chance, C. (2022). Has there been a second AI big Bang? Forbes. Available at: <https://www.forbes.com/sites/calumchance/2022/10/18/has-there-been-a-second-ai-big-bang/?sh=7f7860fb6f74> (Accessed January 1, 2023).
- Chavez, H., Chavez-Arias, B., Contreras-Rosas, S., Alvarez-Rodríguez, J. M., and Raymundo, C. (2023). Artificial neural network model to predict student performance using non-personal information. *Front. Educ.* 8:1106679. doi: 10.3389/feduc.2023.1106679
- Chen, S. (2023). ChatGPT-5 rumored to launch at the end of 2023, will it achieve AGI? [blog post]. Available at: <https://www.gizmochina.com/2021/04/01/chatgpt-5-rumored-to-launch-at-the-end-of-2023-will-it-achieve-agi/> (Accessed April 2, 2023).
- Clarizia, F., Colace, F., Lombardi, M., Pascale, F., and Santaniello, D. (2018). Chatbot: an education support system for students. In: A. Castiglione, F. Pop, M. Ficco and F. Palmieri (Eds) *Cyberspace safety and security*, 11161. Springer: Cham
- ColdFusion. (2022). It's time to pay attention to A.I. (ChatGPT and beyond) [video]. YouTube. Available at: Retrieved from <https://www.youtube.com/watch?v=0uQqMxXoNVs> (Accessed January 1, 2023).
- Crossplag (n.d.). AI content detector. Available at: <https://crossplag.com/ai-content-detector/> (Accessed April 8, 2023).
- Cunningham-Nelson, S., Boles, W., Trouton, L., and Margerison, E. (2019). A review of Chatbots in education: practical steps forward. Proceedings of the AAEE2019 conference Brisbane, Australia. Available at: https://aaee.net.au/wp-content/uploads/2020/07/AAEE2019_Annual_Conference_paper_184.pdf
- Dastin, J., Hu, K., and Dave, P. (2022). Exclusive: ChatGPT owner OpenAI projects \$1 billion in revenue by 2024. Reuters. Retrieved from (<https://www.reuters.com/business/chatgpt-owner-openai-projects-1-billion-revenue-by-2024-sources-2022-12-15/>).
- Durall, E., and Kapros, E. (2020). Co-design for a competency self-assessment Chatbot and survey in science education. In: P. Zaphiris and A. Ioannou (Eds) *Learning and collaboration technologies. Human and technology ecosystems*, Springer: Cham, 12206
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., et al. (2023). "So what if ChatGPT wrote it?" multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *Int. J. Inf. Manag.* 71:102642. doi: 10.1016/j.ijinfomgt.2023.102642
- Elecrow. (2023). Introducing ChatGPT4: the next generation of language models. Elecrow. Available at: <https://www.elecrow.com/blog/introducing-chatgpt4-the-next-generation%C2%A0of%C2%A0language-models.html> (Accessed March 3, 2023).
- Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). GPTs are GPTs: an early look at the labor market impact potential of large language models. Available at: <https://arxiv.org/abs/2023.10130v4> (Accessed April 1, 2023).
- Essel, H., Vlachopoulos, D., Tachie-Menson, A., Johnson, E., and Baah, P. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *Int. J. Educ. Technol. High. Educ.* 19, 1–19. doi: 10.1186/s41239-022-00362-6
- Falke, T., Ribeiro, L., Utama, P., Dagan, I., and Gurevych, I. (2019). Ranking generated summaries by correctness: an interesting but challenging application for natural language inference. In Proceedings of the 57th annual meeting of the Association for Computational Linguistics, 2214–2220
- Felix, C. (2020). "The role of the teacher and AI in education" in *International perspectives on the role of Technology in Humanizing Higher Education Innovations in higher education teaching and learning*. eds. E. Sengupta, P. Blessinger and M. S. Makhanya, vol. 33 (Bingley: Emerald Publishing Limited), 33–48.
- Felten, E., Raj, M., and Seamans, R. (2023). How will language modelers like ChatGPT affect occupations and industries? *arXiv preprint arXiv:2303.01157v2*. doi: 10.48550/arXiv.2303.01157
- Fitzpatrick, K., Darcy, A., and Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR Ment. Health* 4:e19. doi: 10.2196/mental.7785
- Fryer, L., Ainley, M., Thompson, A., Gibson, A., and Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: an experimental comparison of chatbot and human task partners. *Comput. Human Behav.* 75, 461–468. doi: 10.1016/j.chb.2017.05.045

- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L., and Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: a randomized controlled trial. *JMIR Ment. Health* 5:e64. doi: 10.2196/mental.9782
- Gaceta, La (2023). Prohiben usar la ChatGPT en varias universidades. Available at: <https://www.lagaceta.com.ar/nota/978320/sociedad/prohiben-usar-chatgpt-varias-universidades.html> (Accessed February 5, 2023).
- Gao, C., Howard, F., Markov, N., Dyer, E., Ramesh, S., Luo, Y., et al. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *bioRxiv preprint* 2022.12.23.521610. doi: 10.1101/2022.12.23.521610
- Georgescu, A. (2018). Chatbots for education - trends, benefits, and challenges. *Int. Sci. Conf. eLearn. Softw. Educ.* 2, 195–200. doi: 10.12753/2066-026X-18-097
- GESTION (2023). Los primeros trabajos perdidos en Estados Unidos por la Inteligencia Artificial ChatGPT. Available at: <https://gestion.pe/economia/management-emprego/chatgpt-los-primeros-trabajos-perdidos-en-estados-unidos-por-la-implementacion-de-la-inteligencia-artificial-nnda-nnlt-noticia/> (Accessed March 3, 2023).
- Ghosal, S. (2023). ChatGPT on characteristic mode analysis. *TechRxiv Preprint*. 70, 1008–1019. doi: 10.36227/techrxiv.21900342.v1
- Goldman, S. (2022). OpenAI CEO admits ChatGPT risks. What now? | the AI beat. VentureBeat. Available at: <https://venturebeat.com/ai/openai-ceo-admits-chatgpt-risks-what-now-the-ai-beat/> (Accessed January 1, 2023).
- GPT, Chat (2022). In Wikipedia. Available at: <https://en.wikipedia.org/wiki/ChatGPT> (Accessed January 1, 2023).
- GPT-2 Output Detector Demo (2022). Hugging Face, Inc. Available at: <https://huggingface.co/openai-detector/> (Accessed January 3, 2023).
- Hatzius, J., Briggs, J., Kodnani, D., and Pierdomenico, G. (2023). The potentially large effects of artificial intelligence on economic growth. Available at: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf (Accessed April 1, 2023).
- Hien, H., Cuong, P.-N., Nam, L., Nhung, H., and Thang, L. (2018). Intelligent assistants in higher-education environments: the fit-ebot, a chatbot for administrative and learning support. Proceedings of the 9th international symposium on info and communication technology, 69–76
- Hu, J. (2020). Online test proctoring claims to prevent cheating. But at what cost? SLATE. Available at: <https://slate.com/technology/2020/10/online-proctoring-proctoru-proctoru-cheating-research.html> (Accessed January 3, 2023).
- Hu, K. (2023). ChatGPT sets record for fastest-growing user base - analyst note. Reuters. Available at: <https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/> (Accessed February 5, 2023).
- Insights, Oxford. (2022). Government AI readiness index 2022. Available at: https://static1.squarespace.com/static/58b2e92c1e5b6c828058484e/t/639b495cc6b59c620c3ecde5/1671121299433/Government_AI_Readiness_2022_FV.pdf
- Islam, A., Kader, F., and Shin, S. (2018). BSSQS: a Blockchain-based smart and secured scheme for question sharing in the Smart education system. *arXiv preprint arXiv:1812.03917v1 [cs.CR]*. doi: 10.48550/arXiv.1812.03917
- Ismail, M., and Ade-Ibajola, A. (2019). Lecturer's apprentice: a chatbot for assisting novice programmers. IEEE Xplore, 2019 international multidisciplinary information technology and engineering conference (IMITEC), 1–8
- Jeblick, K., Schachtner, B., Dextl, J., Mittermeier, A., Stüber, A., Topalis, J., et al. (2022). ChatGPT makes medicine easy to swallow: an exploratory case study on simplified radiology reports. *arXiv preprint arXiv:2212.14882v1 [cs.CL]*. doi: 10.48550/arXiv.2212.14882
- Kamalov, F., Suliman, H., and Calonge, D. (2021). Machine learning-based approach to exam cheating detection. *PLoS One* 16:e0254340. doi: 10.1371/journal.pone.0254340
- Kasapalu, R., Prieto, L. P., Ley, T., and Chejara, P. (2022). Teacher artificial intelligence-supported pedagogical actions in collaborative learning coregulation: a wizard-of-Oz study. *Front Educ.* 7, 1–15. doi: 10.3389/feduc.2022.736194
- Kerly, A., and Bull, S. (2006). The potential for chatbots in negotiated learner modeling: a wizard-of-oz study. *Lect. Notes Comput. Sci.* 4053, 443–452. doi: 10.1007/11774303_44
- Kim, S., Lee, J., and Gweon, G. (2019). "Comparing data from Chatbot and web surveys" in *Proceedings of the 2019 CHI conference on human factors in computing systems - CHI*, vol. 19, 1–12.
- King, M. R. chatGPT (2023). A conversation on artificial intelligence, Chatbots, and plagiarism in higher education. *Cell. Mol. Bioeng.* 16, 1–2. doi: 10.1007/s12195-022-00754-8
- Klos, M., Escoredo, M., Joerin, A., Lemos, V., Rauws, M., and Bunge, E. (2021). Artificial intelligence-based Chatbot for anxiety and depression in university students: pilot randomized controlled trial. *JMIR Formative Res.* 5:e20678. doi: 10.2196/20678
- L'Heureux, A., Grolinger, K., and Capretz, M. A. M. (2022). Transformer-based model for electrical load forecasting. *Eng.* 15:4993. doi: 10.3390/en15144993
- Leavy, S. (2018). Gender bias in artificial intelligence: the need for diversity and gender theory in machine learning. In *Proceedings of the 1st international workshop on gender equality in software engineering (GE '18)*. New York, NY, USA, 14–16
- Lee, Y. F., Hwang, G. J., and Chen, P. Y. (2022). Impacts of an AI chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educ. Technol. Res. Dev.* 70, 1843–1865. doi: 10.1007/s11423-022-10142-8
- Libert, K. (2023). Only 34% of educators support the decision to ban ChatGPT. EINPRESSWIRE, January 16, 2023. Available at: <https://www.einpresswire.com/article/611597643/only-34-of-educators-support-the-decision-to-ban-chatgpt> (Accessed January 22, 2023).
- Lopez-Lira, A., and Tang, Y. (2023). Can ChatGPT forecast stock Price movements? Return Predictability and Large Language Models. SSRN Preprint. Available at: <http://dx.doi.org/10.2139/ssrn.4412788>
- Lukpat, A. (2023). ChatGPT banned in new York City public schools over concerns about cheating, learning development. Available at: <https://www.wsj.com/articles/chatgpt-banned-in-new-york-city-public-schools-over-concerns-about-cheating-learning-development-11673024059> (Accessed at January 8, 2023).
- Lund, B., and Agbaji, D. (2023). Information literacy, data literacy, privacy literacy, and ChatGPT: Technology literacies align with perspectives on emerging technology adoption within communities (January 14, 2023). SSRN Preprint. Available at: <http://dx.doi.org/10.2139/ssrn.4324580>
- Mabunda, K., and Ade-Ibajola, A. (2019). Pathbot: an intelligent chatbot for guiding visitors and locating venues. IEEE Xplore, 2019 6th international conference on soft computing & machine intelligence (ISCMi), 160–168
- Maynez, J., Narayan, S., Bohnet, B., and McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. In *Proceedings of the 58th annual meeting of the Association for Computational Linguistics*, 1906–1919
- McCallum, S. (2023). ChatGPT banned in Italy over privacy concerns. BBC News. Available at: <https://www.bbc.com/news/technology-60929651> (Accessed April 13, 2023).
- Meckler, L., and Verma, P. (2022). Teachers are on alert for inevitable cheating after the release of ChatGPT. The Washington Post. Available at: <https://www.washingtonpost.com/education/2022/12/28/chatbot-cheating-ai-chatbotgpt-teachers/> (Accessed January 2, 2023).
- Miller, F., Katz, J., and Gans, R. (2018). AI x I = AI2: the OD imperative to add inclusion to the algorithms of artificial intelligence. *OD Practitioner* 5, 6–12.
- Murtarellia, G., Gregory, A., and Romentia, S. (2021). A conversation-based perspective for shaping ethical human-machine interactions: the particular challenge of chatbots. *J. Bus. Res.* 129, 927–935. doi: 10.1016/j.jbusres.2020.09.018
- Norris, M. (2019). University online cheating -- how to mitigate the damage. *Res. High. Educ. J.* 37, 1–20.
- Nunamaker, J., Derrick, D., Elkins, A., Burgoon, J., and Patton, M. (2011). Embodied conversational agent-based kiosk for automated interviewing. *J. Manag. Inf. Syst.* 28, 17–48. doi: 10.2753/MIS0742-1222280102
- O'Donovan, M. A., McCallion, P., McCarron, M., Lynch, L., Mannan, H., and Byrne, E. (2019). A narrative synthesis scoping review of life course domains within health service utilisation frameworks [version 1; peer review: 2 approved]. *HRB Open Res.* 2, 1–18. doi: 10.12688/hrbopenres.12900.1
- Okonkwo, C. W., and Ade-Ibajola, A. (2021). Chatbots applications in education: a systematic review. *Comput. Educ.* 2:100033. doi: 10.1016/j.caeai.2021.100033
- OpenAI (2022). Playground. Available at: www.OpenAI.com and <https://beta.openai.com/playground>
- OpenAI (2023). AI text classifier [website]. Available at: <https://beta.openai.com/ai-text-classifier> (Accessed April 8, 2023).
- OpenAI (2023). GPT-4. Available at: <https://openai.com/research/gpt-4> (Accessed April 9, 2023).
- OpenAI (2023). GPT-4 technical report. *arXiv preprint arXiv:2303.08774*. doi: 10.48550/arXiv.2303.08774
- Ordóñez, V., Dunn, T., and Noll, E. (2023). OpenAI CEO Sam Altman says AI will reshape society, acknowledges risks: 'a little bit scared of this'. Available at: <https://abcnews.go.com/Technology/openai-ceo-sam-altman-ai-reshape-society-acknowledges/story?id=76971838> (Accessed April 1, 2023).
- Originality. AI – plagiarism checker and AI detector (2023). Available at: <https://originality.ai/about> (Accessed January 3, 2023).
- Ortiz, S. (2022). What is ChatGPT, and why does it matter? Here's what you need to know. ZDNET. Available at: <https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-what-you-need-to-know/> (Accessed January 1, 2023).
- Page, L., and Gehlbach, H. (2017). How an artificially intelligent virtual assistant helps students navigate the road to college. *AERA Open* 3:233285841774922. doi: 10.1177/2332858417749220
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. doi: 10.1136/bmj.n71
- Patel, S. B., and Lam, K. (2023). ChatGPT: the future of discharge summaries? *Lancet Digit. Health* 5, e107–e108. doi: 10.1016/S2589-7500(23)00021-3

- Pavlik, J. (2023). Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education. *J. Mass Commun. Educ.* 78, 84–93. doi: 10.1177/10776958221149577
- Perry, N., Srivastava, M., Kumar, D., and Boneh, D. (2022). Do users write more insecure code with AI assistants? *arXiv preprint arXiv:2211.03622v2 [cs.CR]*. doi: 10.48550/arXiv.2211.03622
- Peters, M. D. J., Godfrey, C. M., Khalil, H., McInerney, P., Parker, D., and Soares, C. B. (2015). Guidance for conducting systematic scoping reviews. *Int. J. of Evid. Based Healthc.* 13, 141–1461. doi: 10.1097/XEB.0000000000000050
- Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., et al. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evi. Synthesis* 18, 2119–2126. doi: 10.1111/2474-1015.200167
- Pham, X., Pham, T., Nguyen, Q., Nguyen, T., and Cao, T. (2018). Chatbot as an intelligent personal assistant for mobile language learning. ICEEL 2018: Proceedings of the 2018 2nd international conference on education and E-learning, 16–21
- Pickard, M., Schuetzler, R., Valacich, J., and Wood, D. (2017). Next-generation accounting interviewing: a comparison of human and embodied conversational agents (ECAs) as interviewers. *SSRN Electron. J.* April, 1–21. doi: 10.2139/ssrn.2959693
- Point, Check (2023). OPWNAI: Cybercriminals starting to use CHATGPT. Available at: <https://research.checkpoint.com/2023/opwnai-cybercriminals-starting-to-use-chatgpt/> (Accessed January 18, 2023).
- Qin, C., Zhang, A., Zhang, Z., Chen, J., Yasunaga, M., and Yang, D. (2023). Is ChatGPT a general-purpose natural language processing task solver? *arXiv preprint arXiv:2302.06476*. doi: 10.48550/arXiv.2302.06476
- Rahim, N. I. M., Iahad, N. A., Yusof, A. F., and Al-Sharafi, M. A. (2022). AI chatbots adoption model for higher-education institutions: a hybrid PLS-SEM-neural network modelling approach. *Sustainability* 14:12726. doi: 10.3390/su141912726
- Ramos-Pérez, E., Alonso-González, P., and Núñez-Velázquez, J. (2021). Multi-transformer: a new neural network-based architecture for forecasting S&P Volatility. *Math* 9:1794. doi: 10.3390/math9151794
- Ranoliya, B., Raghuwanshi, N., and Singh, S. (2017). Chatbot for university-related faqs. IEEE Xplore. 2017 international conference on advances in computing, communications, and informatics (ICACCI), 1525–1530
- Reis-Marques, C., Figueiredo, R., and Neto, M. (2021). Applications of Blockchain technology to higher education arena: a bibliometric analysis. *Eur. J. Investig. Health Psychol. Educ.* 11, 1406–1421. doi: 10.3390/ejihpe11040101
- Reza, S., Campos, M., Machado, J., and Tavares, J. (2022). A multi-head attention-based transformer model for traffic flow forecasting with a comparative analysis of recurrent neural networks. *Expert Syst. Appl.* 202:117275. doi: 10.1016/j.eswa.2022.117275
- Rodchua, S., Yaaiodom-boakye, G., and Woolsey, R. (2011). Student verification system for online assessments: bolstering quality and integrity of distance learning. *J. Technol. Manag. Appl. Eng.* 27, 1–8.
- Roose, K. (2022). The brilliance and weirdness of ChatGPT. The New York Times. Available at: <https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html> (Accessed January 1, 2023).
- Ruipérez-Valiente, J., Muñoz-Merino, P., Alexandron, G., and Pritchard, D. (2017). Using machine learning to detect 'multiple-account' cheating and analyze the influence of student and problem features. *IEEE Trans. Learn. Technol.* 12, 112–122. doi: 10.1109/TLT.2017.2784420
- Ryan-Mosley, T. (2023). Five big takeaways from Europe's AI act. MIT Technology Review. Available at: <https://www.technologyreview.com/2023/06/19/1075063/five-big-takeaways-from-europes-ai-act/> (Accessed June 25, 2023).
- Sandu, N., and Gide, E. (2022). Adoption of AI-Chatbots to enhance student learning experience in higher education in India. IEEE Xplore. 2019 18th Int. Conf. On information technology based higher education and training (ITHET)
- Sangalli, V., Martínez-Muñoz, G., and Cañabate, E. (2020). "Identifying cheating users in online courses" in *IEEE Xplore. 2020 IEEE global engineering education conference (EDUCON)*, 1168–1175.
- Shankland, S. (2022). Why Everyone's obsessed with ChatGPT, a mind-blowing AI Chatbot. CNET. Available at: <https://www.cnet.com/tech/computing/why-everyones-obsessed-with-chatgpt-a-mind-blowing-ai-chatbot/> (Accessed January 1, 2023).
- Shorey, S., Ang, E., Yap, J., Ng, E., Lau, S., and Chui, C. (2019). A virtual counseling application using artificial intelligence for communication skills training in nursing education: development study. *J. Med. Internet Res.* 21:e14658. doi: 10.2196/14658
- Shrivastava, R. (2022). Teachers fear ChatGPT will make cheating easier than ever. Forbes. Available at: <https://www.forbes.com/sites/rashishrivastava/2022/12/12/teachers-fear-chatgpt-will-make-cheating-easier-than-ever/?sh=5b0ab2ac1eef> (Accessed on January 2, 2023).
- Shuster, K., Poff, S., Chen, M., Kiela, D., and Weston, J. (2021). Retrieval augmentation reduces hallucination in conversation. *arXiv preprint arXiv:2104.07567v1 [cs.CL]*. doi: 10.48550/arXiv.2104.07567
- Sinha, S., Basak, S., Dey, Y., and Mondal, A. (2020). An educational Chatbot for answering queries. In: J. Mandal and D. Bhattacharya (Eds) *Emerging Technology in Modelling and Graphics*, 937. Springer: Singapore
- Slepankova, M. (2021). Possibilities of artificial intelligence in education: an assessment of the role of AI Chatbots as a communication medium in higher education (dissertation). Available at: <http://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-108427> (Accessed January 1, 2023).
- Smith, C. (2023). GPT-5 might make ChatGPT indistinguishable from a human. BGR. Available at: <https://bgr.com/tech/gpt-5-might-make-chatgpt-indistinguishable-from-a-human/> (Accessed April 1, 2023).
- Sobania, D., Briesch, M., Hanna, C., and Petke, J. (2023). An analysis of the automatic bug fixing performance of ChatGPT. *arXiv preprint arXiv:2301.08653*. doi: 10.48550/arXiv.2301.08653
- Soper, T. (2023). Seattle public schools bans ChatGPT; district 'requires original thought and work from students'. GeekWire. Available at: <https://www.geekwire.com/2023/seattle-public-schools-bans-chatgpt-district-requires-original-thought-and-work-from-students/> (Accessed January 1, 2023).
- Stanley, J., Brink, R., Valiton, A., Bostic, T., and Scollan, R. (2022). "Chatbot accessibility guidance: a review and way forward" in *Proceedings of sixth international congress on information and communication technology*. eds. X. S. Yang, S. Sherratt, N. Dey and A. Joshi, vol. 216 (Singapore: Springer)
- Susnjak, T. (2022). ChatGPT: the end of online exam integrity? *arXiv preprint arXiv:2212.09292v1 [cs.AI]*. doi: 10.48550/arXiv.2212.09292
- Tallyn, E., Fried, H., Gianni, R., Isard, A., and Speed, C. (2018). The Ethnobot: gathering ethnographies in the age of IoT. Conference on human factors in computing systems - proceedings
- Tate, K. (2021). Artificial intelligence will destroy 'laptop class' workers. The Hill. Available at: <https://thehill.com/opinion/technology/541331-artificial-intelligence-will-destroy-laptop-class-workers> (Accessed April 2, 2023).
- Taulli, T. (2019). *Artificial intelligence basics: A non-technical introduction*. 1st Edn. CA: Apress Berkeley.
- Tegos, S., Demetriadis, S., and Karakostas, A. (2015). Promoting academically productive talk with conversational agent interventions in collaborative learning settings. *Comput. Educ.* 87, 309–325. doi: 10.1016/j.compedu.2015.07.014
- The Washington Post - Editorial Board (2022). Opinion. We asked an AI bot hundreds of questions. Here's what we learned. Available at: <https://www.washingtonpost.com/opinions/interactive/2022/chatgpt-questions-predictions-ethics/> (Accessed January 2, 2023).
- Thorp, H. (2023). ChatGPT is fun, but not an author. *Science* 379:313. doi: 10.1126/science.adg7879
- Treize, K., Ryan, T., Barba, P., and Kennedy, G. (2019). Detecting contract cheating using learning analytics. *J. Learn. Anal.* 6, 90–104. doi: 10.18608/jla.2019.63.11
- Tsai, C., and Wu, J. (2022). "A Blockchain-based fair and transparent homework grading system for online education" in *Principles and practice of blockchains*. eds. K. Daimi, I. Dionysiou and N. El Madhoun (Cham: Springer)
- Vallés-Pérez, I., Soria-Olivas, E., Martínez-Sober, M., Serrano-López, A., Gómez-Sanchis, J., and Mateo, F. (2022). Approaching sales forecasting using recurrent neural networks and transformers. *Expert Syst. Applica.* 201:116993. doi: 10.1016/j.eswa.2022.116993
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., and Polosukhin, I. (2017). Attention is all you need. *arXiv preprint arXiv:1706.03762 [cs.CL]* 31st Conf. On neural Inf. Process. Systems (NIPS 2017), Long Beach, CA, USA
- Walters, L. (2021). Universities expand controversial AI monitoring of students taking online exams. Stuff. Available at: <https://www.stuff.co.nz/national/education/126908715/universities-expand-controversial-ai-monitoring-of-students-taking-online-exams> (Accessed January 4, 2023).
- Wang, Y., Petrina, S., and Feng, F. (2017). VILLAGE—virtual immersive language learning and gaming environment: immersion and presence. *Br. J. Educ. Technol.* 48, 431–450. doi: 10.1111/bjet.12388
- Writer. (n.d.). AI content detector | GPT-3 | GPT-4 | ChatGPT. Available at: <https://writer.com/ai-content-detector/> (Accessed April 8, 2023).
- Wu, N., Green, B., Ben, X., and O'Banion, S. (2020). Deep transformer models for time series forecasting: the influenza prevalence case. *arXiv preprint arXiv:2001.08317v1 [cs.LG]*. doi: 10.48550/arXiv.2001.08317
- Xiao, Z., Zhou, M. X., Chen, W., Yang, H., and Chi, C. (2020). If I hear you correctly: building and evaluating interview Chatbots with active listening skills. In CHI '20: Proceedings of the 2020 CHI conference on human factors in computing systems 1–14)
- Xiao, Z., Zhou, M., Liao, Q., Mark, G., Chi, C., Chen, W., et al. (2020). Tell me about yourself: using an AI-powered Chatbot to conduct conversational surveys. *ACM Trans. Comput.-Hum. Interact.* 27, 1–37. doi: 10.1145/3381804
- Zhai, X. (2023). ChatGPT for next generation science learning. *SSRN Preprint. (January 20, 2023)*. doi: 10.2139/ssrn.4331313
- Zhao, J., Song, T., and Sun, Y. (2020). Apihelper: helping junior android programmers learn API usage. *IAENG Int. J. Comput. Sci.* 47, 1–16.



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Bob Edmison,
Virginia Tech, United States
Rosanna Yuen-Yan Chan,
The Chinese University of Hong Kong, China

*CORRESPONDENCE

Michael Yee
✉ myee@ll.mit.edu

RECEIVED 30 June 2023

ACCEPTED 29 August 2023

PUBLISHED 15 September 2023

CITATION

Yee M, Roy A, Perdue M, Cuevas C, Quigley K,
Bell A, Rungta A and Miyagawa S (2023)
AI-assisted analysis of content, structure, and
sentiment in MOOC discussion forums.
Front. Educ. 8:1250846.
doi: 10.3389/feduc.2023.1250846

COPYRIGHT

© 2023 Yee, Roy, Perdue, Cuevas, Quigley, Bell,
Rungta and Miyagawa. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
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.

AI-assisted analysis of content, structure, and sentiment in MOOC discussion forums

Michael Yee^{1*}, Anindya Roy², Meghan Perdue²,
Consuelo Cuevas¹, Keegan Quigley¹, Ana Bell³, Ahaan Rungta²
and Shigeru Miyagawa⁴

¹Artificial Intelligence Technology Group, MIT Lincoln Laboratory, Lexington, MA, United States, ²Open Learning, Massachusetts Institute of Technology, Cambridge, MA, United States, ³Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, United States, ⁴Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA, United States

Discussion forums are a key component of online learning platforms, allowing learners to ask for help, provide help to others, and connect with others in the learning community. Analyzing patterns of forum usage and their association with course outcomes can provide valuable insight into how learners actually use discussion forums, and suggest strategies for shaping forum dynamics to improve learner experiences and outcomes. However, the fine-grained coding of forum posts required for this kind of analysis is a manually intensive process that can be challenging for large datasets, e.g., those that result from popular MOOCs. To address this issue, we propose an AI-assisted labeling process that uses advanced natural language processing techniques to train machine learning models capable of labeling a large dataset while minimizing human annotation effort. We fine-tune pretrained transformer-based deep learning models on category, structure, and emotion classification tasks. The transformer-based models outperform a more traditional baseline that uses support vector machines and a bag-of-words input representation. The transformer-based models also perform better when we augment the input features for an individual post with additional context from the post's thread (e.g., the thread title). We validate model quality through a combination of internal performance metrics, human auditing, and common-sense checks. For our Python MOOC dataset, we find that annotating approximately 1% of the forum posts achieves performance levels that are reliable for downstream analysis. Using labels from the validated AI models, we investigate the association of learner and course attributes with thread resolution and various forms of forum participation. We find significant differences in how learners of different age groups, gender, and course outcome status ask for help, provide help, and make posts with emotional (positive or negative) sentiment.

KEYWORDS

MOOCs, discussion forums, forum posts, natural language processing, text classification, machine learning, transformers, artificial intelligence

1. Introduction

Massive Open Online Courses (MOOCs) are tremendous educational resources for learners seeking to educate themselves or gain new skills. A key component of MOOCs is the discussion forums, a place for learners to engage in conversation, provide and receive help, and establish a learning community. Although studies have not found evidence that

strong social networks are formed in MOOC discussion forums (Gillani and Eynon, 2014; Boroujeni et al., 2017; Wise et al., 2017), they do create learning communities where learners can get support (Poquet and Dawson, 2015). Additionally, studies have shown that giving and receiving help from student peers increases learning gains (Topping, 2005; Yamarik, 2007). Others have shown that engagement in discussion forums in MOOCs is linked to learner retention in the course (Houston et al., 2017; Poquet et al., 2018). However, analyzing discussion forum use is challenging due to the vast amount of unstructured data and the complexity of the interactions. Artificial Intelligence (AI) offers an opportunity to overcome these difficulties and provide valuable insights into the learning process within discussion forums. This research will address the following questions:

- RQ1: Can we use AI algorithms to tag forum posts along category, structure, and emotion dimensions as reliably as human coders? If the answer to RQ1 is yes, we can use AI-generated tags to answer the following research questions:
- RQ2: How are the tags distributed in the forums and within threads? How are the tags related to each other?
- RQ3: How are learner attributes and course attributes associated with the forum participants? Specifically, how are these attributes associated with whether a thread started by a learner got resolved, and what is the likelihood of such a learner posting a comment of a certain type?

2. Related work

2.1. Online education and forums

Researchers have sought to identify patterns in discussion forum usage that could shed light on which learners are using the forums, and how they are using them. Many studies have shown that there are variations of forum usage, with some learners using the forums more than others, though the most consistent group is active learners interested in completing the course (Huang et al., 2014; Almatrafi and Johri, 2018; Moreno-Marcos et al., 2018). Research into forum usage over time has found that the number of learners using the forum and overall quantity of posts diminish through the course run (Brinton et al., 2014). However, other studies (Wong et al., 2015; Galikyan et al., 2021) found that learners used progressively higher cognitive levels as they advanced throughout the course.

Forums support a broad range of social activities in the course, such as small talk, questions about the logistics, help-seeking and help-giving behaviors, and content-based discussion. Studies have investigated the impact of participating in content vs. non-content-based discussion threads, and found that engagement in the forums was positively correlated with course performance, regardless of post type (Wise et al., 2017; Wise and Cui, 2018). Boroujeni et al. (2017) found that the forums play a particularly useful role in content-triggering discussions, especially for help-seekers. Some research has been conducted to seek to unpack the degree to which the forums are effective for learners who need additional support with the content (Kim and Kang, 2014). Yang et al. (2015) looked specifically at the impact of unresolved confusion or help-seeking

on learners. They found that expressing confusion in the forums was negatively correlated with course retention, though this was mediated by receiving support and resolving the confusion.

Relatedly, there has been interest in whether the emotional sentiment of a learner's post could predict their retention in the course. Ezen-Can et al. (2015) sought to understand the relationship between the emotional sentiment of a learner's posts and their persistence in the course by creating a sentiment score for each learner based on all their posts. They looked at three distinct MOOCs and found that the patterns differed for each course, but for the computer science course, a significant emotional response (positive or negative) was correlated with a higher dropout rate. Wen et al. (2014) sought to model learner dropout rates based on a sentiment analysis of learner discussion forums posts and usage type and found that positive and actively engaged learners were most likely to complete the course, followed by active and negative.

MOOC forum posts have also been analyzed in connection to participants' demographic attributes, such as gender and age. In a study conducted by Swinnerton et al. (2017), it was discovered that older learners are more likely to post comments. Another study by Huang et al. (2014) compared forum superposters to ordinary participants in terms of age and gender. The findings revealed that superposters tend to be older than the average forum users, and there is a small but statistically significant over-representation of women among superposters, while there are generally more male forum participants. Gender differences in MOOC forum posts have been explored by John and Meinel (2020) by analyzing the types of questions asked by men and women, as well as the categories of responses these questions generated. Although they did not find statistical significance between the question types and gender, they observed that male learners tended to participate more in longer discussions.

2.2. Automated forum post classification

Studying forum data requires characterizing individual posts along one or more dimensions, such as topic, activity, degree of confusion, and sentiment. This data enrichment process can require a significant investment of time and/or money especially if the dataset is large. However, each dimension can be formulated as a multi-class text classification problem, and natural language processing (NLP) techniques can be used to help automate or semi-automate the labeling process.

Early applications of machine learning to text classification involved representing textual content as a bag of words (or longer n-grams), with terms weighted by their term frequency-inverse document frequency (TF-IDF) weights, and training shallow machine learning models such as support vector machine (SVM), logistic regression, and random forest classifiers (Schütze et al., 2008).

Like other domains, NLP was revolutionized by deep learning (LeCun et al., 2015). Pretrained word embeddings

such as word2vec (Mikolov et al., 2013a,b) and GloVe (Pennington et al., 2014) introduced a powerful alternative to the bag-of-words representation, and a variety of neural network architectures such as multilayer perceptrons (MLPs), recurrent neural networks (RNNs), convolutional neural networks (CNNs), and transformer models have achieved state of the art performance on many NLP tasks [see Minaee et al. (2021) and Li et al. (2022) for overviews of text classification].

The same progression from traditional (shallow) machine learning methods to deep learning has taken place in the domain of forum post classification (Ahmad et al., 2022). Since the present work explores three distinct classification tasks (category, structure, and emotion), the use of transformer-based models, and strategies for reducing human annotation burden, we highlight relevant work in these areas.

The classification task studied in Ntourmas et al. (2019, 2021) is nearly identical to our category task; Ntourmas et al. (2019) classified starting posts as content-related, logistics-related, or other using an SVM with bag-of-words representation, while Ntourmas et al. (2021) applied a decision tree to a TF-IDF weighted bag-of-words representation along with additional features derived from a seeded topic modeling technique. Sentiment analysis is a common forum post classification task studied in, e.g., Bakharia (2016), Chen et al. (2019); Clavié and Gal (2019), Li et al. (2019), and Capuano et al. (2021). Sentiment is also one of the six dimensions (question, opinion, answer, sentiment, urgency, and confusion) included in the Stanford MOOCPosts dataset (Agrawal et al., 2015), slices of which have been used by many works, e.g., Bakharia (2016), Chen et al. (2019), Clavié and Gal (2019), Guo et al. (2019), Sun et al. (2019), and Alrajhi et al. (2020). For automatic analysis of thread structure, Sun et al. (2016) classified posts within a thread according to dialogue acts (question, answer, resolution, reproduction, other) and also whether one post contains an immediate follow-up discussion of another using conditional random fields (CRFs). Joksimović et al. (2019) used an unsupervised approach combining hidden Markov models (HMMs) and Latent Dirichlet Allocation (LDA) to discover and analyze speech act categories. Fisher et al. (2015) trained an HMM-like latent variable model using weak supervision to classify whether one post is a direct response to another.

Recent works spanning numerous forum post classification tasks have incorporated pretrained transformers such as BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al., 2018), e.g., Clavié and Gal (2019), Li et al. (2019), Sha et al. (2021), Zou et al. (2021), and Lee et al. (2022). BERT is a promising technique since it is often possible to fine-tune models pretrained on large unlabeled datasets using a limited amount of labeled task-specific data.

Other works addressing the problem of annotation burden include Ntourmas et al. (2021), which investigated how many weeks of labeled data are required to perform reliably when applied to the remaining weeks of the MOOC and finds that maximum performance is achieved by week 3, and Chen et al. (2019), which developed a semi-supervised learning approach based on co-training and found that it outperforms traditional and deep

learning baselines while starting with only 30% of the dataset labeled.

3. Methodology

Our overall approach is to create a training set by manually annotating a relatively small subset of the full forum post dataset, train AI models on the labeled data, apply the best-performing models to the unlabeled remainder of the data, and finally use the fully labeled dataset (containing AI-generated predictions plus a small fraction of human-supplied labels) to answer downstream education-related research questions.

To answer RQ1, we implement the following AI-assisted labeling process:

- **Step 1:** Develop initial coding guidelines
- **Step 2:** Annotate new subset of unlabeled posts
- **Step 3:** Measure inter-annotator agreement and adjudicate labels when annotators disagree
- **Step 4:** Refine coding guidelines and return to Step 2 (if necessary due to low agreement)
- **Step 5:** Train AI models on annotated data
- **Step 6:** Evaluate model performance with internal metrics and comparison to human performance, and return to Step 2 to increase training set size (if necessary)
- **Step 7:** Apply best models to unlabeled data
- **Step 8:** Audit subset of model predictions and return to Step 2 to increase training set size (if necessary)

Assuming the model predictions pass the final human audit, the fully labeled dataset generated in Step 7 is then used to answer RQ2 and RQ3. We use exploratory data analysis techniques at both the post and thread level to answer RQ2: using Jaccard similarity to measure co-occurrence of tags from the three dimensions (category, structure, and emotion), and aggregating posts within threads to analyze how tag distributions vary with thread length. For RQ3, we use logistic regression to investigate the association of forum behavior with learner and course attributes.

4. MOOC dataset

For this work we studied 11 course instances of two sequential introductory Python MOOCs (referred to as Python-1 and Python-2): six of these eleven courses were Python-1, and the remaining five were Python-2. The courses ran between Spring 2018 and Summer 2021—which includes a total of five course instances that operated during the COVID-19 pandemic. Since these courses were all offered online, they did not undergo any change in logistics during the pandemic, and their content remained the same over this period. All the courses were 9 weeks long and instructor-paced, with defined start and end dates. Approximately 387,000 learners enrolled for these 11 courses, about 6% of whom paid for full access and the opportunity to earn a certificate (“verified learners”), and ~53% of the verified learners went on to earn a certificate. The learners along with their instructors and community teaching assistants (TAs) generated ~82000 posts. Additionally, we have

detailed records of how the learners interacted with various other course materials such as problems and videos, but we do not take that into account for the purpose of this work. For many of these learners we have voluntarily supplied demographic information such as their age, gender, and level of education, and we include that in our work as a representative sample. In addition, we could infer the country where most learners accessed the courses, and note the country's economic category (e.g., high-income country, lower-middle-income country, etc.). In this work, we use learners' verification and certification status, as well as when a particular course ran (i.e., before or during the pandemic) where relevant. For more information on the demographic details of the learners, we refer our readers to the authors' previous work in Roy et al. (2022) and Yee et al. (2022).

5. Annotation process

To create the training dataset for AI model development, our team members ultimately labeled a collection of 950 forum posts randomly sampled from the 11 Python MOOCs, with approximately an equal share coming from each course. Each team member labeled an initial set of 50 posts. The initial labels selected in the project were based on previous research conducted in this area analyzing forum posts (Brinton et al., 2014; Boroujeni et al., 2017; Wise et al., 2017; Galikyan et al., 2021) and optimized for the research questions. The labels were finalized by reviewing sample codes with all coders present and discussing how the labels would apply to the data. Once the group came to a consensus on the application, the labels and definitions were drafted into coding guidelines that were distributed to all coders. The labels for each task were:

- **Category:** Logistics, Content, Emotional/Commentary
- **Structure:** Question, Suggestion/Explanation, Follow Up/Follow Up Question, Resolution, Comment/Response
- **Emotion:** Positive, Negative, Neutral

Sentiment analysis has been studied with various formulations ranging from binary classification (positive/negative) to multi-point scales (Pang and Lee, 2008; Zhang et al., 2018). We use three labels (positive/negative/neutral) to support our downstream analysis.

After labeling, we held an adjudication session where we discussed disagreements, decided on adjudicated (consensus) answers, and refined the coding guidelines. When two or more coders disagreed on a label, a group of three or more coders reviewed the data together and came to a consensus on the correct code to use, and any changes in the interpretation of labeling was detailed in the coding guidelines. The team labeled three additional sets of posts (of size 250, 300, and 350) with two team members annotating each post, followed by additional adjudication sessions. The additional rounds were required to improve either coding understanding and consistency or current model performance (Steps 4 and 6). The final coding guidance is given in Table 1.

We computed Krippendorff's alpha (Krippendorff, 1970) to assess consistency across annotators (inter-annotator agreement) and the quality of the coding guidelines. For the Category,

Structure, and Emotion tasks, we achieved alphas of 0.632, 0.634, and 0.532, respectively. These values are lower than what is generally considered to be very good agreement ($\alpha \geq 0.8$). However, these final values span multiple labeling rounds and evolving annotation team membership. Since we adjudicated all disagreements among annotators for the entire training set, we considered the final labels to be high quality and sufficient for model training and downstream analysis.

In Figure 1, we present the relative percentages of the classes within the three tasks as coded by the human annotators, as well as those predicted by the AI-assisted labeling process (only the subset of the posts made by the learners alone is presented here). The full predicted dataset, which includes posts from instructors and TAs, is closer to the distributions of the manually annotated set. Note that the distributions of the classes within each task revealed moderate to severe class imbalance, e.g., negative posts only made up 5% of the final annotated dataset. This was a key challenge in developing accurate models.

6. Model design and performance

In this section, we describe data preprocessing approaches, model architectures, and training hyperparameters that we explored to develop models capable of accurately labeling the Category, Structure, and Emotion dimensions for unlabeled forum posts.

6.1. Data preprocessing

To address the class imbalance and limited size of the labeled dataset, multiple dataset preprocessing steps were tested. First, contextual information about each forum post was added as input to the model. Discussions in forums on the edX platform take place in threads with the following structure: thread title, initial post, zero or more response posts, and zero or more comments on each response post. For each forum post in the dataset, we reconstructed the containing thread (using forum database tables persisted by edX) and extracted the following thread context features:

- `thread_title`
- `post_type`: initial_post, response_post, or comment
- `num_responses`: number of responses in thread
- `response_position`: position in response list for the response containing this post (if post is response or comment)
- `num_response_comments`: number of comments on the response containing this post (if post is response or comment)
- `comment_position`: position in comment list for the response containing this comment (if post is comment)
- `original_poster`: whether post is authored by the same author as thread's initial post

To aid our annotation process, we condensed these features into a short textual form. For example, a post with context "Response 1/2; Num Comments 7" means it was the first response out of two and it received a total of seven comments.

TABLE 1 Coding guidance for the Category, Structure, and Emotion forum post classification tasks.

Task	Code	Definition
Category	Logistics	Post relates to logistics of using the platform/accessing materials in the course. Typically noting that a problem set isn't working, there is a typo in the materials, their submission didn't go through for some reason, or they are asking for an extension.
Category	Content	Post relates to the content of the course itself. Typically asking a question about the material, clarifying understanding, asking for help with a problem, requesting additional materials to understand the concept, remarking on the videos/materials, etc.
Category	Emotional/Commentary	Post conveys the learners feelings about the class, problem, or experience. Often can be commenting on the pace/difficulty of the course, encouraging others in the class to keep going, etc.
Structure	Question	Post is a question—specifically seeking help for an issue in the course. Often this could be asking for help due to a logistics issue, or help with understanding the content and completing the assignments.
Structure	Suggestion/Explanation	Post is a comment to another learners question post, giving suggestions for how to solve their issue, or trying to explain what the OP doesn't understand.
Structure	Follow Up/Follow Up Question	Post is a follow up to a suggestion/explanation post, but not a resolution. Typically learners would make a follow up post if they attempted the suggestion/explanation and it didn't work, or they want to clarify something.
Structure	Resolution	Post indicates that the question has been resolved. Typically noting that either the suggestion/explanation was sufficient and successful, or noting that it wasn't successful but the learner has given up on the problem or issue.
Structure	Comment/Response	Post is outside the question/suggestion/ follow up/resolution stream. Often general or specific comments about observations, thoughts, feelings.
Emotion	Positive	Post is overwhelmingly positive in sentiment.
Emotion	Negative	Post is overwhelmingly negative in sentiment.
Emotion	Neutral	Post is overwhelmingly neutral in sentiment.

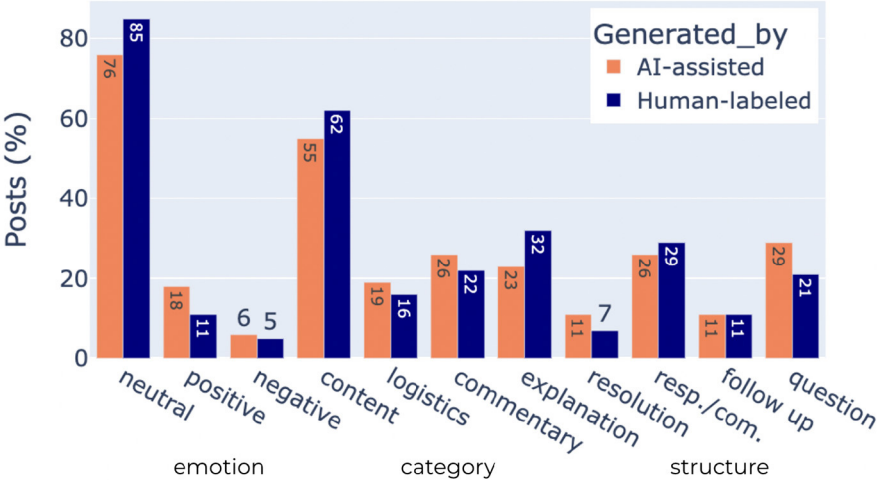


FIGURE 1 Distribution (%) of different classes under each task labeled by AI and human annotators. Classes under each task sum up to 100%. The percentages are obtained from 950 posts labeled by human coders, and from >61,000 learner posts labeled by the trained AI model.

We tested adding this information as input to the model in multiple ways: by concatenating the textual context information to the forum post body; and by constructing a short vector of context features, processing these with a multilayer perceptron (MLP), and concatenating the context embedding with the body embedding before classification.

Because of extreme class imbalance, especially in the case of Negative emotional content, we looked for ways to augment our labeled dataset. Two common methods employed in NLP

tasks for dataset augmentation are backtranslation and synonym replacement, both of which we tested. In backtranslation, monolingual data is auto-translated to a foreign language and back to the original language, leading to slight variations in wording but preserving sentiment and content (Sennrich et al., 2015). Synonym replacement replaces random words in the post with their synonyms, with some rules guiding which words can or cannot be replaced (Niu and Bansal, 2018).

Because backtranslation augmentation creates just a single alternative text body from a forum post and due to computational constraints, we created three augmented text bodies for each labeled forum post, and could randomly draw from the four bodies during training. These augmented bodies (two backtranslation, one synonym replacement) were created using the NLPAug Python package (Ma, 2019), with backtranslation performed through two intermediary languages, German and Russian, using Facebook FAIR's pretrained WMT19 News Translation models (Ng et al., 2019). During training, we tested *no augmentation* (none), random backtranslation (backtranslation), and random backtranslation/synonym replacement (random). For backtranslation and random augmentations, the augmentation was applied to the forum post with a probability of 0.5.

Beyond adding augmentations and additional context vectors, we also tested using weighted random sampling to draw samples from the training dataset for each training batch. This effectively evened the number of training examples from each class. Together with augmentations, we hypothesized that weighted random sampling might lead to higher accuracy across all classes, including those that suffered from severe class imbalance.

6.2. Model architecture

We used a deep learning model consisting of a pretrained BERT encoder (Vaswani et al., 2017; Devlin et al., 2018), and a small MLP classification head that operated on the BERT embeddings. Our classification head consisted of a linear layer with hidden dimension size $h = 128$, ReLU activation, dropout regularization with probability 0.1, and another linear layer projecting from $h \rightarrow \text{num_classes}$. If a context vector was provided to the model, we processed this input vector with an MLP, consisting of a linear layer with hidden dimension 32, ReLU, dropout with probability 0.1, a second linear layer with hidden dimension 32, ReLU, dropout ($p = 0.1$), and a linear projection to a context embedding vector of size 16. This context embedding vector was then appended to the BERT embedding before classification by the classification head.

We tested two pretrained BERT models from Devlin et al. (2018), bert-base-cased ($L = 12$, $H = 768$, Total Params = 110M) and bert-large-cased ($L = 24$, $H = 1024$, Total Params = 340M). We also tested the BERT-tweet-eval-emotion model from HuggingFace, pretrained on the tweet_eval dataset (Rosenthal et al., 2017; Barbieri et al., 2020; Schmid, 2021), for its performance on our emotion classification task.

For our loss function, we experimented with cross entropy loss (CE) and soft cross entropy loss (SCE) based on estimated class probabilities. For CE, we used the unanimous consensus labels or adjudicated labels (from resolved disagreements) as our targets (see Section 5). For SCE, we estimated class probabilities from the annotation and adjudication process by normalizing the individual annotator labels as "votes." The adjudicated labels were also included as a vote when available. This produced estimates of class likelihood between $[0, 1]$, which were used as soft targets for our model. We trained all models for 10 epochs, and used an

TABLE 2 Best model hyperparameters (architecture) and performance for the Structure, Category, and Emotion tasks. Mean accuracy and macro F1 scores across the five test sets from cross-validation are reported, along with standard deviations.

	Structure	Category	Emotion
Augmentation	none	random	backtranslation
Loss function	sce	sce	ce
BERT model	large	large	tweet-eval
Batch size	8	8	16
Context	textual*	vector	none
Accuracy	0.76 ± 0.03	0.82 ± 0.02	0.87 ± 0.03
Macro F1	0.73 ± 0.03	0.77 ± 0.03	0.67 ± 0.08

For Context, textual* indicates textual context format, without a weighted random sampler.

AdamW optimizer ($LR = 2e-5$, $\epsilon = 1e-8$) with linear warmup (2 epochs) and linear decay (8 epochs).

While BERT and other deep learning transformer-based models have dominated the NLP domain in recent years, more traditional ML and NLP tools like SVMs continue to work sufficiently well for certain tasks. We compared our BERT models to SVM "bag-of-words" approaches, establishing a baseline performance level for the three tasks. To construct these models, we used a count vectorizer to transform each text input into a matrix of token counts, transformed this matrix using TF-IDF weighting, and finally fit an SVM to these transformed features and class labels. The SVM used a maximum of 1000 iterations, hinge loss with L2 penalty, and $\alpha = 1e-3$. We tested the SVM with each of our dataset preprocessing options, including augmentations, weighted random sampling, and appending the context and thread title to the post body before input to the model.

6.3. Model experiments

To assess the performance of each model architecture on the Category, Structure, and Emotion tasks, we followed a five-fold cross-validation procedure in our experiments. In this procedure, we split the data into five equally sized folds, using four-folds for training the model, and one-fold as a "holdout set" for evaluating the model. Since this holdout set has not been used during training of the model, predictions on its elements can be used to estimate the performance of the model on the remaining unlabeled data. For each model architecture, we trained five distinct models during five-fold cross-validation, with each model trained with a different fold held out during training. Using this procedure, we obtained five independent estimates of the performance of the model's architecture. Aggregating these estimates (Nadeau and Bengio, 1999), we obtained the mean performance and variance estimate across the entire labeled dataset, and could compare model architectures. Additionally, cross-validation methods allowed us to use ensemble methods to predict labels during inference.

For the deep learning experiments with BERT, we use early stopping to guard against overfitting to the training set (Prechelt, 2012). After removing a holdout fold from the training data during cross-validation, we perform a further split for deep learning

TABLE 3 Model comparison for the Emotion task.

Model	Augmentation	BERT model	Loss	Thread context	Accuracy	Macro F1
Annotator A	—	—	—	—	0.919	0.823
Annotator B	—	—	—	—	0.780	0.735
SVM	random	—	—	textual	0.820 ± 0.022	0.505 ± 0.057
BERT	backtranslation	BERT-tweet-eval-emotion	ce	none	0.873 ± 0.026	0.671 ± 0.081
BERT	none	BERT-tweet-eval-emotion	ce	none	0.860 ± 0.035	0.613 ± 0.095
BERT	backtranslation	bert-base-cased	ce	none	0.862 ± 0.032	0.588 ± 0.071
BERT	backtranslation	BERT-tweet-eval-emotion	ce	textual	0.876 ± 0.015	0.640 ± 0.067
BERT	backtranslation	BERT-tweet-eval-emotion	sce	none	0.885 ± 0.021	0.653 ± 0.078

TABLE 4 Model comparison for the Structure task.

Model	Augmentation	BERT model	Loss	Thread context	Accuracy	Macro F1
Annotator A	—	—	—	—	0.924	0.905
Annotator B	—	—	—	—	0.780	0.641
SVM	random	—	—	textual	0.596 ± 0.050	0.538 ± 0.049
BERT	none	bert-large-cased	sce	textual*	0.758 ± 0.033	0.731 ± 0.030
BERT	random	bert-large-cased	sce	textual*	0.749 ± 0.025	0.711 ± 0.029
BERT	none	bert-base-cased	sce	textual*	0.689 ± 0.041	0.625 ± 0.059
BERT	none	bert-large-cased	sce	textual	0.734 ± 0.023	0.704 ± 0.021
BERT	none	bert-large-cased	ce	textual*	0.738 ± 0.013	0.670 ± 0.052

TABLE 5 Model comparison for the Category task.

Model	Augmentation	BERT model	Loss	Thread context	Accuracy	Macro F1
Annotator A	—	—	—	—	0.939	0.920
Annotator B	—	—	—	—	0.819	0.744
SVM	none	—	—	textual	0.742 ± 0.027	0.679 ± 0.031
BERT	random	bert-large-cased	sce	vector	0.818 ± 0.018	0.768 ± 0.026
BERT	none	bert-large-cased	sce	vector	0.787 ± 0.029	0.730 ± 0.051
BERT	random	bert-base-cased	sce	vector	0.789 ± 0.035	0.735 ± 0.044
BERT	random	bert-large-cased	sce	none	0.783 ± 0.036	0.728 ± 0.042
BERT	random	bert-large-cased	ce	vector	0.779 ± 0.060	0.729 ± 0.055

In Tables 3–5, the model with the best mean Macro F1 score (bolded) was selected as the best model. Annotator Accuracy and Macro F1 is a comparison of individual annotators' labels and the adjudicated labels (for the subset of the dataset that each annotator labeled). The annotators with the best and worst Macro F1 scores are shown. The adjudication process was partially dependent on the individual annotators' labels, so the annotator performance should be considered a *high* target for model performance. For Thread context, textual* indicates textual context format, without a weighted random sampler.

experiments, reserving 10% of the training data as the validation dataset. This validation data is also withheld from training, but after every training epoch, labels for the validation dataset are predicted by the model and the loss is computed. The model from the training epoch with the lowest validation loss is selected as the best model for that fold, and its expected performance is computed using the holdout fold (functioning as the test set).

We conduct a grid search for each of the Category, Structure, and Emotion tasks over various parameters

used to construct a model architecture, finding the best-performing model architecture after training and testing using the five-fold cross-validation scheme. For the grid search over deep learning models using BERT, we sweep over augmentation (none, backtranslation, random), loss function (cross entropy, soft cross entropy), pretrained BERT model (bert-base-cased/base, bert-large-cased/large, as well as BERT-tweet-eval-emotion/tweet for the Emotion

task only), and batch size (8, 16). For each of the gridded parameters, we also experiment with four different configurations of contextual information inputs: information encoded into a six-element context vector, processed with a small MLP, and concatenated to BERT embedding (vector); contextual information concatenated with forum post in a structured way before input to the BERT model (textual); textual without a weighted random sampler to even the number of examples from each class during training (textual*); and no contextual information added (none). This search produced a series of 144 Emotion and 96 Structure/Category model architectures to compare (see Section 6.4).

We also conduct a grid search for the SVM-based models, sweeping over thread context (textual, as above, or none), weighted class sampling (weighted, unweighted), and text augmentation (none, backtranslation, random), producing 12 candidate model architectures for each classification task.

6.4. Model performance

We compare the models from the grid searches in Section 6.3, and report the best model architectures, as determined by mean macro F1 score across five distinct holdout sets from cross-validation (Table 2). We use macro F1 score to select our top models because it provides a better estimate of performance for highly imbalanced classes, like the Negative class in Emotion classification.

This class imbalance is likely a driving factor behind the relatively poor macro F1 score for the Emotion task. Many of the ablation experiments that we tried were targeted to improve this performance, including augmentations and loss function. In Tables 3–5, we report results obtained when a single hyperparameter from the grid search is changed in the best model architecture (bolded) for each of the three tasks. We compare this series of BERT models to the best SVM model, and provide baseline human performance levels from the annotation procedure for comparison. Annotator accuracy/macro F1 is computed on the subset of labeled examples that the annotator contributed a label for, and individual annotator labels are compared to the adjudicated labels (considered truth values). In this way, the annotator predictions are *not independent* from the truth values, and as such, their accuracy and macro F1 scores are considered *upper bounds* on true annotator performance. We report the performance of two of the seven annotators, with Annotators A and B attaining the highest and lowest macro F1 scores among annotators, respectively.

As seen in the ablation tables, the BERT-based models outperform the SVM models. However, the performance of the top BERT-based model is not statistically better than many of the ablation experiments, given the reported standard deviation from the cross-validation test sets. For the Category and Structure tasks, the best-performing models outperform Annotator B's macro F1 score, and thus fall within the spread of human annotator performance. However, this was not the case for best model on the Emotion task, and we suggest a few possible reasons for the task being more difficult.

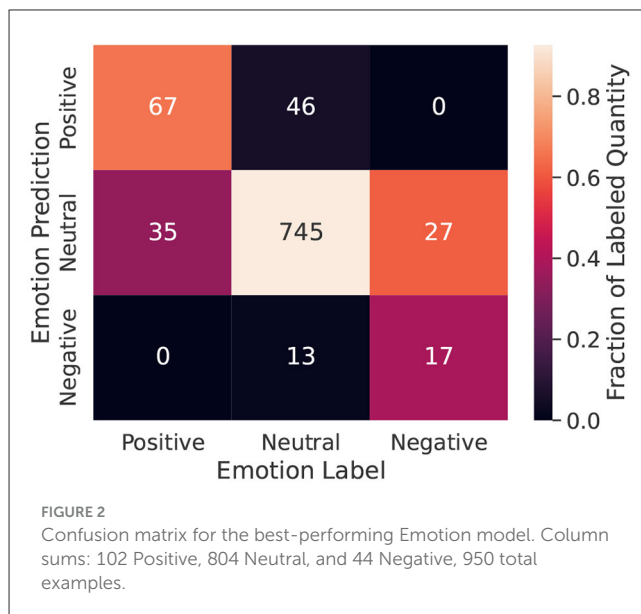
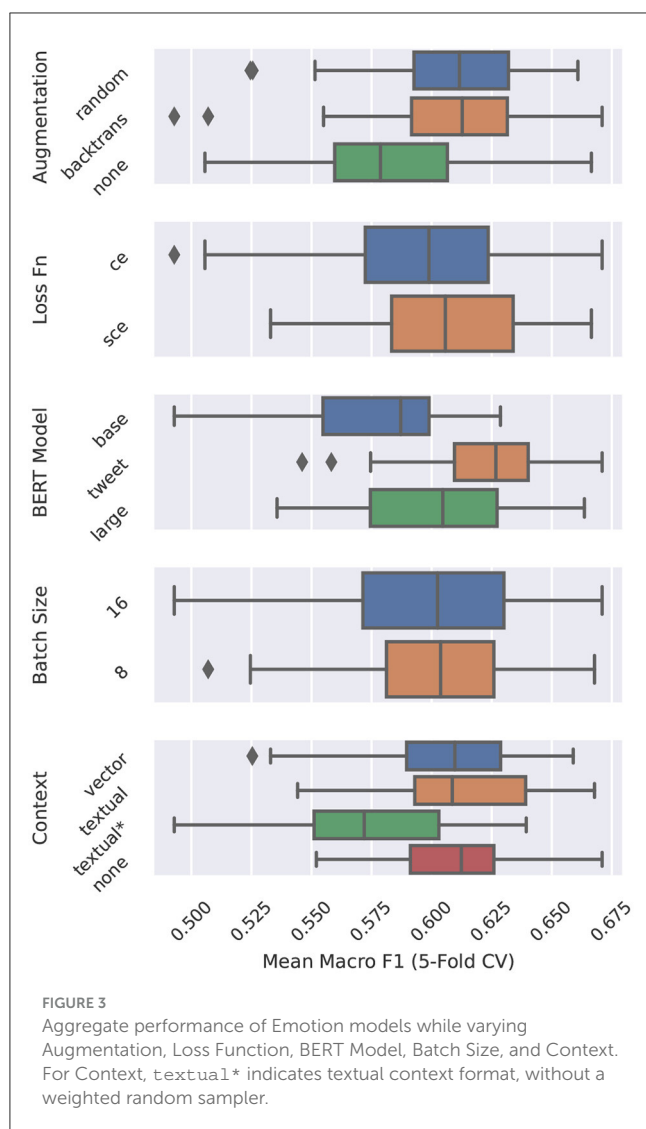


Figure 2 shows the confusion matrix as a heatmap for the best-performing Emotion model. The column sums represent the class distribution in our labeled dataset (also seen in Figure 1), with evident class imbalance. The biggest difficulty appears to be misclassifying Negative emotion posts as Neutral, which is understandable given the Negative class had the smallest number of examples across any class for any task in the 950 post training dataset. Despite these misclassifications, our auditing process (see Section 7) found 97% agreement for the Emotion labels after adjudication, perhaps indicating that many of these misclassified posts were “toss-ups” between Negative and Neutral, and that coding guidelines during initial labeling could have been clearer.

Figure 3 further examines the effect of each hyperparameter in the Emotion classification ablation studies, with a grid search conducted over 144 sets of model hyperparameters. For each hyperparameter value, the macro F1 score is averaged across all models trained with that value. We see, for example, that both random and backtranslation augmentation strategies outperform no augmentation on average. We also see that weighted sampling is very important for the Emotion task since textual*—the only Context hyperparameter setting without weighted sampling—performs worst.

Even though the macro F1 score of the best-performing Emotion model was lower than the score of Annotator B, possibly due to the class imbalance and coding issues outlined above, we considered the model's performance adequate for the Emotion task because the human annotator score is an upper bound (and considered a high target).

Since the focus of our work was to meet or exceed the performance of human annotators, enabling large-scale forum post classification, we didn't perform direct comparisons with alternative modeling approaches (e.g., different sets of features and model architectures). However, our model performance was on par with results from related works conducted on other MOOC forum datasets. For example, for category classification, Ntourmas et al. (2019) report an accuracy of 0.69 for an SVM model trained



on an introductory Python course, and [Ntourmas et al. \(2021\)](#) achieve an accuracy of 0.64 on this course using the first two weeks of posts as training data for a decision tree classifier, and then evaluating the model on data from future weeks (three through six). Our best BERT-based model achieved an accuracy of 0.82 using 1% of the dataset, although we sample training data from all weeks of the MOOCs. For the emotion task, [Clavié and Gal \(2019\)](#) report accuracy of 89.78 for their EduBERT model on the StanfordMOOC sentiment dataset (where the original 7-point scale has been converted to binary with a score of 4 or above considered positive). Our best BERT model achieves 87.26 accuracy on a three-class formulation of the problem for our dataset. Finally, [Sun et al. \(2016\)](#) report an accuracy of 0.576 for a CRF applied to a 12-class dialogue act classification task on an edX MOOC dataset. We achieve an accuracy of 0.76 for our structure task, although our formulation has fewer classes (five).

Having achieved model performance that seemed on par with human annotators, we used the best models for each task to classify the remaining unlabeled data (approximately 81,000 posts). For each task, instead of training a new model on the full training set

of 950 posts using the best hyperparameter setting (from [Table 2](#)), we created an ensemble of the five models trained during the five-fold cross-validation process associated with the best-performing setting. Ensembles have been shown to improve both accuracy and uncertainty calibration assuming enough diversity across ensemble members ([Lakshminarayanan et al., 2017](#)). Each of the five models trained during cross-validation was trained on a slightly different (though overlapping) subset of the training examples and had its non-pretrained weights initialized with different random seeds. Model predictions were then available for auditing (Step 8 in [Section 3](#)).

7. Auditing model outputs

Once the best models were identified, the model outputs were audited by team members for veracity. A team member randomly sampled 250 posts that had not been previously hand-coded. The team member blindly labeled the posts, then compared the hand-labeled posts with the best model label predictions (95% agreement for Category, 79% agreement for Structure, and 89% agreement for Emotion). Then the team-member labels and model labels were compared and adjudicated. After the adjudication session, there was a 96% agreement for Category, 87% for Structure, and 97% for Emotion. The improvements in Structure and Emotion were due to human error, which was detected in the adjudication process.

8. Model applications

With forum comments tagged and validated, we explored various forum characteristics and learner interactions to answer the research questions presented in [Section 1](#). These questions serve as examples of how our AI models could be applied to understand structures and patterns in MOOC discussion forums. For RQ2, we investigated how the Category, Structure, and Emotion tasks are distributed in the forums and their relationship to one another across single posts as well as in comment threads. In RQ3 we asked how learners engaged with the forums, and if any group-level or course-level attributes emerged in their forum engagements. About a quarter of all posts were made by the Community TAs or instructors, and these posts were excluded from our analysis.

8.1. RQ2: use AI-assisted labeling to determine forum structures

In [Figure 1](#), we presented the relative distribution of various classes under each task. As a next step, we analyzed how the different classes under each task overlapped with the classes from the other tasks. Our chosen metric is Jaccard similarity score, which for two sets A and B , is defined as $|A \cap B|/|A \cup B|$. The result is presented as a heatmap in [Figure 4](#). As expected from the analysis of individual comments, there was a large overlap between the Content category and Neutral emotion (Jaccard score of 0.64). The Jaccard score of posts which fell in the Commentary category and Response/Comment structure was 0.44—the highest of any two classes among the structure and category tasks. The

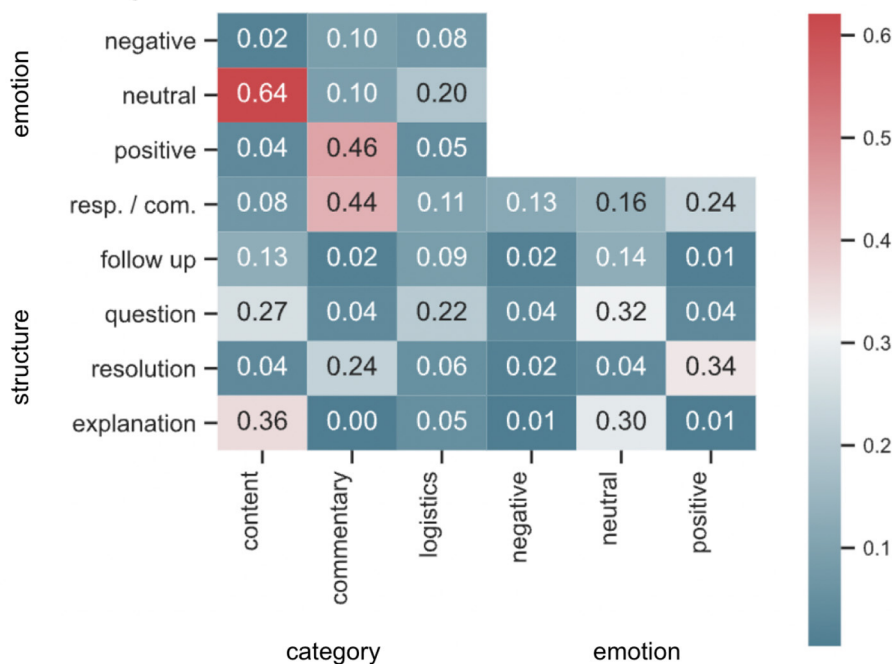


FIGURE 4
Jaccard similarity scores for AI-generated tags.

Commentary category had a higher Jaccard score with Positive emotion (0.46) than Negative emotion (0.10). The Structure tag of type Resolution had the Jaccard score of 0.34 with the Positive Emotion tag: the highest of all Emotion and Structure task pairs. Structure tags of types Question and Suggestion/Explanation were generally more Neutral emotionally, with scores 0.32 and 0.30, respectively. The highest score between the Structure and Category tasks were for posts categorized as Content with structure of type Suggestion/Explanation (0.36). The most common structure for the Logistics-category posts was Question (Jaccard score 0.22). Besides providing insight into the interrelationship among the tags, this analysis serves to further validate the output of our AI models by generating associations that are expected (e.g., the high score between Resolution structure and Positive emotion).

To understand the thread-level organization of forum posts, we aggregated posts for each thread, and determined how the Structure, Category, and Emotion tasks of each thread are distributed (e.g., a thread containing 2 Neutral posts, 1 Positive post, and 1 Negative post would have emotion fractions 0.5, 0.25, and 0.25, respectively). We binned these threads by the number of posts they contained, and presented the fraction of each task class as a function of the number of posts, averaged over the number of threads in a bin. We presented our analysis in Figure 5. There were fewer threads in the bins at the higher end of post counts, and the lines in the plots are more jagged.

We found that threads contained more general Commentary type structures as they grew in length, while the fraction of Questions kept falling. This downward trend of the Question fraction is expected: most threads had one question (and the follow-up questions had their own class). Explanation/Suggestion type

structures showed a peak around 10 posts, followed by a slow descent. For the Emotion task, we found that the Neutral class occupied a smaller fraction in the longer threads, while the Negative and the Positive classes occupied a larger fraction. Similarly, the Commentary category occupied a larger fraction in the longer threads, at the cost of the Content- and Logistics-type categories. As in the previous analysis of Jaccard score among task classes, some of these trends (e.g., the Question fraction vs. the number of posts) returned intuitively expected answers, thus confirming the general validity of our AI-assisted labeling process.

8.2. RQ3: analysis of task associations with learner attributes

In this sub-section we analyze learner interactions with the forums using the AI-generated labels. We investigated two sub-questions: (a) How did the forum participants' demographic attributes and course attributes relate to the likelihood of their questions reaching resolution in a thread? (b) Was there any association between forum participants' attributes and whether they posted a comment in different classes of Structure, Category, and Emotion?

For each of these questions, we defined corresponding binary outcome variables indicating if there is at least one positive case, grouped at the learner level: i.e., to answer part (a) we looked at if a learner started at least one thread that reached a resolution. For part (b), the outcome variable was positive if a learner made at least one post of a certain type. We performed logistic regression with the

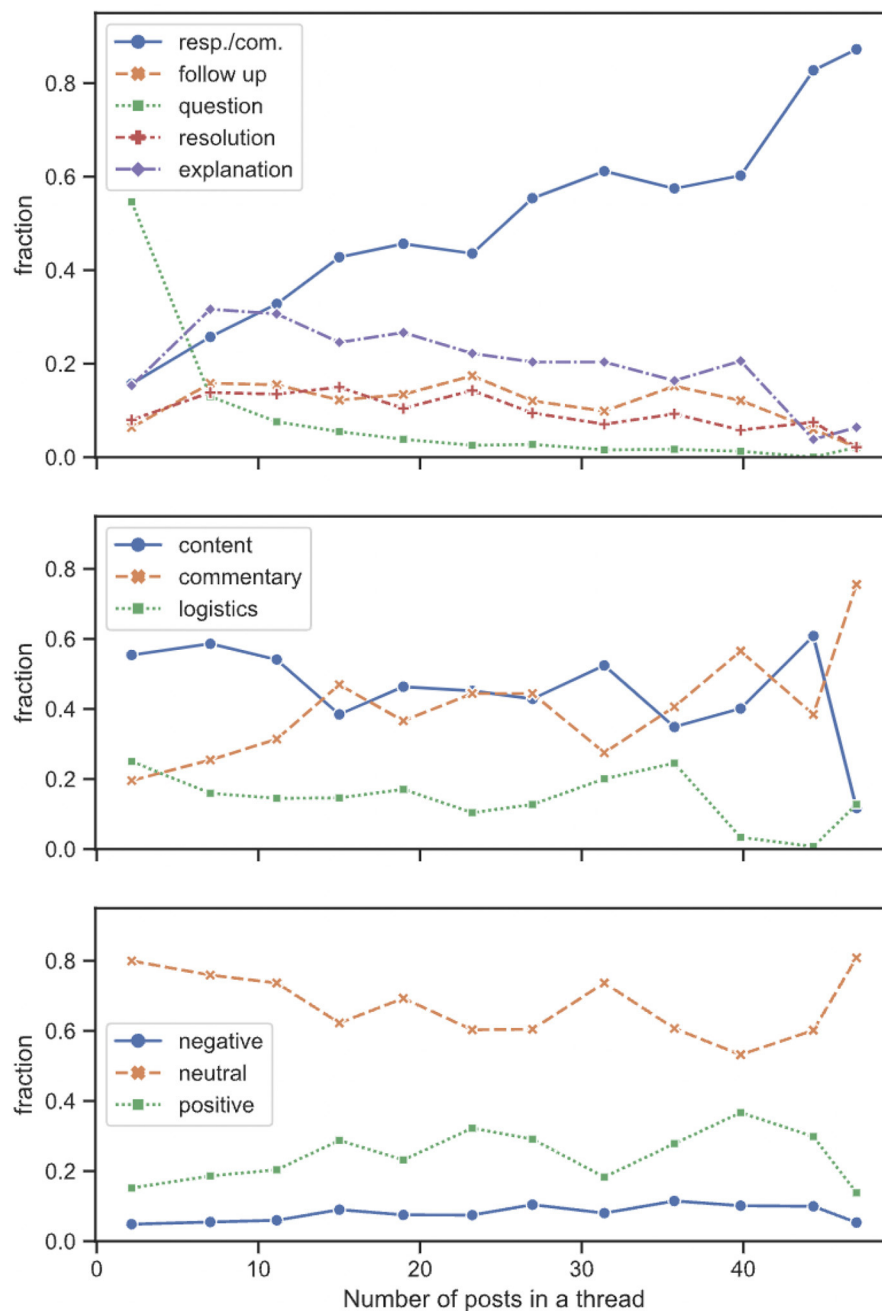


FIGURE 5
Fraction of task classes as functions of average number of posts in threads.

learner attributes and course attributes described below. In both cases, the likelihood would increase if a learner made more posts. To adjust for this, we first studied the number of posts authored by a learner (referred to as `posts_frac`), scaled by the maximum number of posts authored by any learner in our included courses, except the top 1% of posters. We found the top 1% of the forum posters have posted 31–181 posts, and we defined them as outliers, and set the maximum number of posts to 30. The re-scaling process helped us avoid over-weighting this variable when we included it as a covariate in the next analyses. We investigated how the `posts_frac`

variable is associated with the learner attributes using a multiple linear regression model.

Within learner demographics, we considered their gender (male or female), age (below 22, 22–44, over 44 years old), education level (below bachelor's level, bachelors and above), and the economic status of their countries (high-income, upper- and lower-middle-income countries). For the learners' certification/verification status, there were three levels: those who were not verified (ineligible to earn certificates), those who were verified but did not earn certificates (`ver_not_cert`), and those

TABLE 6 Dependence of posts_frac on learner and course attributes. lower_ci and upper_ci refer to the lower and the upper bound of the 95% confidence intervals.

	coef	p-value	lower_ci	upper_ci
Intercept	0.0797	0.0000	0.0700	0.0893
gender_male	0.0017	0.6650	−0.0060	0.0093
age_22_to_44	0.0065	0.1024	−0.0013	0.0143
age_above_44	0.0398	0.0000	0.0296	0.0500
ver_cert	0.0702	0.0000	0.0637	0.0767
ver_not_cert	0.0140	0.0030	0.0048	0.0233
course_2	−0.0015	0.7468	−0.0104	0.0074

who were verified and did earn a certificate (ver_cert). We included the following course attributes: if the learner was enrolled in Python-1 or Python-2, and if the course ran during the COVID-19 pandemic or before it began (pandemic- or pre-pandemic-course). To make our regression analysis robust, we excluded posts where learner gender was recorded as “other,” and if the learner was located in low-income countries, which reduced the learner count by less than 1%. However, such omission prompted by a need for greater statistical clarity reduces our understanding of the learners, and a larger learner sample size in a future study might mitigate this shortcoming.

In all of these regression models, the sample sizes were pretty large, ranging from ~ 3000 to ~ 8500 learners. When working with larger sample sizes, it is common to obtain small *p*-values. To avoid attributing statistical significance where it may not truly exist, we set the threshold *p*-value for significance at 0.0005. However, for the purpose of addressing part (b) of RQ3, we performed multiple tests and imposed a more stringent *p*-value threshold of 0.0001. The *p*-value thresholds are provided here as a heuristic to identify more important results, while the coefficients derived from our regression analyses indicate the effect size within these contexts.

In Table 6, we analyzed how posts_frac depended on the learner/course attributes, via a multiple linear regression model. In the final version of this model, we omitted demographic variables such as the education level, country income category, and whether the course ran during the pandemic or not, based on minimizing the *Akaike Information Score* (AIC). We found the biggest effect to be from the verified learners who went on to earn certificates: they posted about 7% more posts than those who were not verified (and did not earn certificates), after we controlled for other effects. The other significant result is from the learners > 44 years old, who posted about 4% more on average. This result is important, as we find that these two groups (learners > 44 years old and those who earned certificates) have significant results in our other models as well, even after we included the posts_frac as a covariate—meaning that these groups had a direct effect as well as an indirect effect (via posts_frac variable) on the outcome variables. The adjusted R-squared value for this model was ~0.06.

To answer question (a), we first identified threads where the original posts were in the Content or Logistics category with Question-type structure. We aggregated such threads for each learner who started them and counted if at least one thread (for

TABLE 7 Odds ratios, *p*-values, and the associated lower and upper bounds of 95% confidence intervals for the logistic regression model output for the learners whose questions are more likely to get resolved.

	OR	p-value	lower_ci	upper_ci
Intercept	0.2450	0.0000	0.2054	0.2922
gender_male	0.7080	0.0000	0.6025	0.8319
ver_cert	1.4905	0.0000	1.2882	1.7246
ver_not_cert	1.3107	0.0097	1.0677	1.6090
course_2	1.5086	0.0000	1.2552	1.8132
posts_frac	1e+02	0.0000	7e+01	2e+02

a learner) had a Resolution post, as determined by AI labeling. We did not explicitly take into account how many threads a learner started, but included posts_frac as a covariate. We presented the results from our logistic regression modeling for question (a) in Table 7. We chose the simplest model based on the pseudo-R-squared values, and in the final model we excluded learners' education level, country income category, and when the courses ran (i.e., during or before the pandemic). The overall analysis had a $p < 0.0001$. From our results it appeared that the learners who earned certificates and learners in course_2 had almost 50% higher likelihood of reaching a resolution in the threads they started, while male learners had ~30% less odds of reaching a resolution, everything else being equal. The number of posts made by a learner was highly significant and the more a learner posted, the more likely it was that their thread would be resolved. In interpreting these odds ratios we need to remember that posts_frac is scaled to range from 0 to 1. Other factors (including the omitted ones) did not reach the level of significance we set earlier.

To answer question (b), we wanted to determine how learner and course attributes were associated with the likelihood that a learner made at least one post of a certain kind. To achieve this, we aggregated all posts made by a learner in a course, and defined the outcome variable as the likelihood of posting at least one post of a certain class (e.g., Question under the Structure task), as inferred from the AI-generated labels. We included posts_frac as a covariate in this case as well. We limited our analysis to six cases for brevity: Question and Suggestion/Explanation classes from the structure tasks, Positive and Negative classes from the emotion tasks, and Logistics and Content classes from the category tasks. We performed six separate logistic regression analyses: one for each of the above cases. In each of these cases, we aggregated all posts made by a user and noted if they made at least one post of that kind (e.g., negative emotion). For the independent variables, the final models excluded the education level, country income category, and course_pandemic, based on the same selection criteria we used in the last analysis.

In Table 8, we presented the odds ratios and the *p*-value resulting from these six models. With the *p*-value threshold of 0.0001, we found that compared to learners without verification, those who earned certificates were more likely to post at least one comment in all these categories except the structure type Suggestion/Explanation, and in the Content category. The trends were similar for verified learners without certificates, except they

TABLE 8 Odds ratios for included variables in logistic regression models.

	Structure question	Structure explain	Emotion negative	Emotion positive	Category logistics	Category content
Intercept	0.6588*	0.1548*	0.0649*	0.3101*	0.3097*	0.4102*
gender_male	0.8419	1.5624*	0.9093	0.6407*	0.9155	1.1508
age_22_to_44	0.6283*	1.3486*	0.9910	1.2133	0.7758*	1.3021*
age_above_44	0.6028*	1.3503	1.1330	1.0406	0.9493	1.2011
ver_cert	1.4768*	0.8846	1.3175	1.2597*	1.6228*	0.8900
ver_not_cert	1.4636*	0.6418*	1.5801*	1.3372*	1.6137*	0.7388*
course_2	1.2468	1.0231	1.0410	1.0154	0.9937	1.2279
posts_frac	9e+03*	7e+03*	3e+02*	2e+03*	2e+03*	5e+07*

Column names represent the dependent variables for individual regression models. Odds ratios with an asterisk (*) next to them indicate that they are statistically significant at $p \leq 0.0001$ level.

had significantly less likelihood in the Suggestion/Explanation (structure) or in the Content category. Both these groups (verified, with or without certificates) had a much higher likelihood of posting at least one Logistics-type post, and posts with Positive or Negative emotion, than all other groups. Male learners were found to be about 35% less likely to post at least one Positive-emotion post compared to female learners, but there was no significant difference between them when it came to Negative-emotion posts. Other things being equal, men were more likely to post at least one post of Suggestion/Explanation type, and less likely to ask questions. The last trend was similar for age groups 22–44 years old as well as for those 44 and above (compared to learners younger than 22 years old): the former two groups were less likely to post questions, and more likely to post suggestions/explanations. In addition, the learners in the 22–44 years old group were also more likely to post in the Content category, and less likely to post in the Logistics category. As expected, posts_frac strongly increased the likelihood in all six cases. Additionally, we noticed a couple of inversely related trends: the likelihoods of posting a Question and a Suggestion/Explanation (under the Structure tasks) were inversely related, and so were the likelihoods for the Content and the Logistics categories. In Section 9.2, we discuss possible explanations and implications of these findings.

9. Discussion

9.1. Implications of AI-assisted labeling

We successfully applied the proposed AI-assisted labeling process to generate a fully labeled dataset while only requiring our team to annotate approximately 1% of the posts. The AI-generated labels were found to be reliable and formed the foundation of analyses that spanned the category, structure, and emotion dimensions as well as learner demographics and course attributes.

However, the ability to directly apply the models trained on the Python MOOC dataset to other MOOCs is unknown [although Bakharia (2016) and Ntourmas et al. (2019) report poor cross-domain performance in their settings]. Future work could explore the impact of classification task

(Category, Structure, or Emotion) on transferability to other courses.

Even if the models trained on the edX Python courses don't transfer well, we believe this general approach can be extended to non-CS courses on any platform, with appropriate training and testing. There was little direct impact of content on model accuracy, except its homogeneity, and the context added to the posts are replicable from most other forums, making this process platform-independent. One promising future direction of our work would be to evaluate this method for non-CS online courses running on different platforms.

Although we achieved a substantial reduction in annotation effort, making application to new courses feasible, the reduction achieved will likely vary by dataset. Future directions to reduce the burden further include: using active learning (Ren et al., 2021) to identify the most impactful/useful next posts to annotate, using “data programming” techniques (Ratner et al., 2016, 2017) to programmatically generate labels using heuristics, trying other techniques for dealing with class imbalance such as explicitly oversampling likely examples from the minority classes for annotation or applying techniques like SMOTE (Chawla et al., 2002) during model training, using semi-supervised techniques such as self-training as a way to leverage the annotated examples more effectively by pseudo-labeling unlabeled examples, and exploring zero-shot or few-shot learning with large language models (LLMs) such as GPT-3 (Brown et al., 2020). Annotating whole threads at a time (vs. individual posts sampled from disjoint threads) might potentially make the annotation task easier and more reliable (in terms of inter-annotator agreement). Analogously, graph neural networks (Wu et al., 2020) could potentially increase consistency of post labels within threads as well as boost accuracy for the Structure task.

Lastly, recent advances in instruction-tuned LLMs (Ouyang et al., 2022) such as ChatGPT (OpenAI, 2023) suggest the possibility of using an LLM-based agent to draft responses to questions brought up in the forum, which could help improve responsiveness to learners and reduce the burden on course staff. Though a promising future direction, several LLM challenges (OpenAI, 2023; Touvron et al., 2023) would need to be addressed, including hallucination (generating content that is untruthful), producing biased or toxic content, and dealing with new or changing information not present in the training set.

9.2. Implications of our findings from model applications

The improvement in coding performance from using the AI-assisted labeling made it possible to extensively analyze learners' engagement in the discussion forums, and this knowledge could be helpful in multiple ways. When re-designing a course, instructors can preemptively take action to identify and address some recurring issues, such as commonly occurring logistics problems or content-based questions. During an ongoing course with a high volume of forum posts, instructors/TAs may choose to prioritize some posts over others (e.g., the logistics-related posts or those with negative emotions).

Our analysis of ordinary forum participants (i.e., excluding the outliers) show that learners who earned certificates post more, as well as the learners older than 44 years. These results corroborate earlier research findings which showed a correlation between forum posts and course performance (Wen et al., 2014; Houston et al., 2017; Wise and Cui, 2018). On the other hand, we did not observe a strong effect of gender once we controlled for other factors. By bringing together these demographic attributes and course attributes along with the AI-assisted annotations, the present study furthers our understanding of MOOC forum dynamics.

The differences among groups in how learners participated in the forum are often intuitive. For example, verified learners/certificate earners asked more questions in the Logistics category. In MOOCs, these groups of learners often progress more quickly through the course than non-verified learners, and thus these groups are more likely to face issues related to logistics before others. Once these issues are brought to notice, they may get fixed or some workaround is posted in the forum, thus reducing the need for posting Logistics category posts for those attempting the same course component later. On the other hand, the non-verified learners may face more challenges in the content area, and may not be able to find answers on the forums, as the verified learners may not have experienced the same difficulty with content, or they have moved on to a different section of the course. The finding that the verified learners post more comments with positive or negative emotions may be a clue that they are more engaged with the course and more invested in their outcomes, and the joy and frustration of learning they experience are more pronounced than the non-verified learners. This observation was also seen by Wen et al. (2014), who found that both strong positive and negative sentiments were associated with high completion rates for active learners.

From our analysis, we found that verified learners were less likely to post suggestions and explanations. It may be argued that they were better prepared than unverified learners to offer suggestions, but less willing. However, the more likely explanation could be that they may not encounter many other learners seeking help at a point when they are engaged with specific content. If this is indeed true, a course re-design where more advanced learners were encouraged to share their challenges with others is likely to benefit everyone.

From the same analysis, we found systemic differences in how men and women use the forums. For instance, fewer men are likely to post comments with positive emotion than women (while no such difference exists for negative emotion). Whether such

differences make women participants feel less welcome on online forums is not immediately obvious, but is worth exploring further.

While we focused on a handful of applications of AI-assisted labeling, we can extend the same tools to analyze how the forum interactions of individual users change with course progression. We can study the formation of learner groups within forums and the intragroup interaction. This same process could be extended to other courses where a tasks-based description of forums is meaningful, and to other learning platforms where we can find similar context to what we provided to our AI models. We believe that learners who participate in online forums are more motivated to do well in the courses, and the AI-assisted labeling methods to analyze forum participation can provide insight with less human labor, thus enabling course designers and instructors to serve their learners better.

10. Conclusion

This paper sought to explore how learners are engaging with discussion forums in MOOCs to gain support and help each other through the course. We were able to develop a BERT-based model architecture and training approach that supports fine-tuning pretrained models to custom forum post classification tasks using only a small amount of manually generated annotations. We then could use this model to categorize discussion posts to enable an analysis of the data. We showed that different groups of learners often differ significantly in how they interact with the forums: whether they discuss course content or the logistics, the sentiment they show in their posts, and whether they ask questions or offer suggestions. We found that learner attributes of those who start comments threads have strong association with the likelihood that a thread reaches resolution. Future research can build on this work, exploring learner engagement in forums and the impacts of targeted interventions.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: the datasets are not readily available because of potential inclusion of sensitive personal information such as username or email ids in the forum comments, which are difficult to systematically remove. Research access to the data may be available with appropriate authorization and approval from concerned institutional review boards. Requests to access these datasets should be directed to AB, anabell@mit.edu.

Ethics statement

Ethical approval was not required for the study involving human data in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required in accordance with the national legislation and the institutional requirements.

Author contributions

MY, ARo, MP, CC, and SM contributed to the conception and design of the study. AB provided access to the MOOC datasets. MY, ARo, MP, CC, KQ, and ARu annotated data. CC and KQ designed, trained, and analyzed AI models. MP audited model outputs. ARo performed data analysis of model outputs and learner data. MY, ARo, MP, and KQ wrote sections of the manuscript. All authors approved the submitted version.

Funding

Research was sponsored by the United States Air Force Research Laboratory and the Department of the Air Force Artificial Intelligence Accelerator and was accomplished under Cooperative Agreement Number FA8750-19-2-1000.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships

that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Author disclaimer

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Department of the Air Force or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

- Agrawal, A., Venkatraman, J., Leonard, S., and Paepcke, A. (2015). *YouEDU: Addressing Confusion in MOOC Discussion Forums by Recommending Instructional Video Clips*. International Educational Data Mining Society.
- Ahmad, M., Junus, K., and Santoso, H. B. (2022). Automatic content analysis of asynchronous discussion forum transcripts: a systematic literature review. *Educ. Inform. Technol.* 27, 11355–11410. doi: 10.1007/s10639-022-11065-w
- Almatrafi, O., and Johri, A. (2018). Systematic review of discussion forums in massive open online courses (MOOCs). *IEEE Trans. Learn. Technol.* 12, 413–428. doi: 10.1109/TLT.2018.2859304
- Alrajhi, L., Alharbi, K., and Cristea, A. I. (2020). "A multidimensional deep learner model of urgent instructor intervention need in MOOC forum posts," in *Intelligent Tutoring Systems: 16th International Conference, ITS 2020* (Athens: Springer), 226–236.
- Bakharia, A. (2016). "Towards cross-domain MOOC forum post classification," in *Proceedings of the Third 2016 ACM Conference on Learning@ Scale* (New York, NY), 253–256. Available online at: <https://dl.acm.org/doi/abs/10.1145/2876034.2893427>
- Barbieri, F., Camacho-Collados, J., Espinosa-Anke, L., and Neves, L. (2020). "TweetEval: unified benchmark and comparative evaluation for Tweet classification," in *Proceedings of Findings of EMNLP* (Stroudsburg, PA). Available online at: <https://aclanthology.org/2020.findings-emnlp.148/>
- Boroujeni, M. S., Hecking, T., Hoppe, H. U., and Dillenbourg, P. (2017). "Dynamics of MOOC discussion forums," in *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (New York, NY), 128–137. Available online at: <https://dl.acm.org/doi/abs/10.1145/3027385.3027391>
- Brinton, C. G., Chiang, M., Jain, S., Lam, H., Liu, Z., and Wong, F. M. F. (2014). Learning about social learning in MOOCs: from statistical analysis to generative model. *IEEE Trans. Learn. Technol.* 7, 346–359. doi: 10.1109/TLT.2014.2337900
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., et al. (2020). "Language models are few-shot learners," in *Advances in Neural Information Processing Systems 33*, eds H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Red Hook, NY: Curran Associates, Inc.), 1877–1901.
- Capuano, N., Caballé, S., Conesa, J., and Greco, A. (2021). Attention-based hierarchical recurrent neural networks for MOOC forum posts analysis. *J. Ambient Intell. Hum. Comput.* 12, 9977–9989. doi: 10.1007/s12652-020-02747-9
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *J. artif. Intell. Res.* 16, 321–357. doi: 10.1613/jair.953
- Chen, J., Feng, J., Sun, X., and Liu, Y. (2019). Co-training semi-supervised deep learning for sentiment classification of MOOC forum posts. *Symmetry* 12, 8. doi: 10.3390/sym12010008
- Clavié, B., and Gal, K. (2019). EduBERT: pretrained deep language models for learning analytics. *arXiv preprint arXiv:1912.00690*. doi: 10.48550/arXiv.1912.00690
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). BERT: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. doi: 10.48550/arXiv.1810.04805
- Ezen-Can, A., Boyer, K. E., Kellogg, S., and Booth, S. (2015). "Unsupervised modeling for understanding MOOC discussion forums: a learning analytics approach," in *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (New York, NY), 146–150. Available online at: <https://dl.acm.org/doi/abs/10.1145/2723576.2723589>
- Fisher, R., Simmons, R., and Malin-Mayor, C. (2015). "Weakly supervised learning of dialogue structure in MOOC forum threads," in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)* (Miami, FL: IEEE), 624–627. Available online at: <https://ieeexplore.ieee.org/abstract/document/7424387>
- Galikyan, I., Admiraal, W., and Kester, L. (2021). MOOC discussion forums: the interplay of the cognitive and the social. *Comput. Educ.* 165, 104133. doi: 10.1016/j.compedu.2021.104133
- Gillani, N., and Eynon, R. (2014). Communication patterns in massively open online courses. *Internet Higher Educ.* 23, 18–26. doi: 10.1016/j.iheduc.2014.05.004
- Guo, S. X., Sun, X., Wang, S. X., Gao, Y., and Feng, J. (2019). Attention-based character-word hybrid neural networks with semantic and structural information for identifying of urgent posts in MOOC discussion forums. *IEEE Access* 7, 120522–120532. doi: 10.1109/ACCESS.2019.2929211
- Houston, S. L., Brady, K., Narasimham, G., and Fisher, D. (2017). "Pass the idea please: the relationship between network position, direct engagement, and course performance in MOOCs," in *Proceedings of the Fourth 2017 ACM Conference on Learning@ Scale* (New York, NY), 295–298. Available online at: <https://dl.acm.org/doi/abs/10.1145/3051457.3054008>
- Huang, J., Dasgupta, A., Ghosh, A., Manning, J., and Sanders, M. (2014). "Superposter behavior in MOOC forums," in *Proceedings of the First ACM Conference on Learning@ Scale Conference* (New York, NY), 117–126. Available online at: <https://dl.acm.org/doi/abs/10.1145/2556325.2566249>
- John, C., and Meinel, C. (2020). "Learning behavior of men and women in MOOC discussion forums—a case study," in *2020 IEEE Global Engineering Education Conference (EDUCON)* (Porto: IEEE), 300–307. Available online at: <https://ieeexplore.ieee.org/abstract/document/9125322>

- Joksimović, S., Jovanović, J., Kovanović, V., Gašević, D., Milikić, N., Zouaq, A., and Van Staaldin, J. P. (2019). Comprehensive analysis of discussion forum participation: from speech acts to discussion dynamics and course outcomes. *IEEE Trans. Learn. Technol.* 13, 38–51. doi: 10.1109/TLT.2019.2916808
- Kim, J., and Kang, J.-H. (2014). Towards identifying unresolved discussions in student online forums. *Appl. Intell.* 40, 601–612. doi: 10.1007/s10489-013-0481-1
- Krippendorff, K. (1970). Estimating the reliability, systematic error and random error of interval data. *Educ. Psychol. Meas.* 30, 61–70.
- Lakshminarayanan, B., Pritzel, A., and Blundell, C. (2017). “Simple and scalable predictive uncertainty estimation using deep ensembles,” in *Advances in Neural Information Processing Systems 30*, eds I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Red Hook, NY: Curran Associates, Inc.), 6402–6413.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539
- Lee, J., Soleimani, F., Hosmer, J. IV, Soylu, M. Y., Finkelberg, R., and Chatterjee, S. (2022). Predicting cognitive presence in at-scale online learning: MOOC and for-credit online course environments. *Online Learn.* 26, 58–79. doi: 10.24059/olj.v26i1.3060
- Li, Q., Peng, H., Li, J., Xia, C., Yang, R., Sun, L., et al. (2022). A survey on text classification: from traditional to deep learning. *ACM Trans. Intell. Syst. Technol.* 13, 1–41. doi: 10.1145/3495162
- Li, X., Zhang, H., Ouyang, Y., Zhang, X., and Rong, W. (2019). “A shallow bert-cnn model for sentiment analysis on MOOCs comments,” in *2019 IEEE International Conference on Engineering, Technology and Education (TALE)* (Yogyakarta: IEEE), 1–6. Available online at: <https://ieeexplore.ieee.org/abstract/document/9225993>
- Ma, E. (2019). *NLP Augmentation*. Available online at: <https://github.com/makcedward/nlpaug>
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. doi: 10.48550/arXiv.1301.3781
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. (2013b). “Distributed representations of words and phrases and their compositionality,” in *Advances in Neural Information Processing Systems 26*, eds C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Red Hook, NY: Curran Associates, Inc.), 3111–3119.
- Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., and Gao, J. (2021). Deep learning-based text classification: a comprehensive review. *ACM Comput. Surv.* 54, 1–40. doi: 10.1145/3439726
- Moreno-Marcos, P. M., Alario-Hoyos, C., Muñoz-Merino, P. J., Estevez-Ayres, I., and Kloos, C. D. (2018). A learning analytics methodology for understanding social interactions in MOOCs. *IEEE Trans. Learn. Technol.* 12, 442–455. doi: 10.1109/TLT.2018.2883419
- Nadeau, C., and Bengio, Y. (1999). “Inference for the generalization error,” in *Advances in Neural Information Processing Systems*, Vol. 12, eds S.olla, T. Leen, and K. Müller (Cambridge, MA: MIT Press), 307–313.
- Ng, N., Yee, K., Baevski, A., Ott, M., Auli, M., and Edunov, S. (2019). Facebook fair’s WMT19 news translation task submission. *arXiv preprint arXiv:1907.06616*. doi: 10.48550/arXiv.1907.06616
- Niu, T., and Bansal, M. (2018). “Adversarial over-sensitivity and over-stability strategies for dialogue models,” in *The SIGNLL Conference on Computational Natural Language Learning (CoNLL)* (Stroudsburg, PA). Available online at: <https://aclanthology.org/K18-1047/>
- Ntourmas, A., Avouris, N., Daskalaki, S., and Dimitriadis, Y. (2019). “Comparative study of two different MOOC forums posts classifiers: analysis and generalizability issues,” in *2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA)* (Patras: IEEE), 1–8. Available online at: <https://ieeexplore.ieee.org/abstract/document/8900682>
- Ntourmas, A., Dimitriadis, Y., Daskalaki, S., and Avouris, N. (2021). “Classification of discussions in MOOC forums: an incremental modeling approach,” in *Proceedings of the Eighth ACM Conference on Learning@ Scale* (New York, NY), 183–194. Available online at: <https://dl.acm.org/doi/abs/10.1145/3430895.3460137>
- OpenAI (2023). Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*. doi: 10.48550/arXiv.2303.08774
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., et al. (2022). “Training language models to follow instructions with human feedback,” in *Advances in Neural Information Processing Systems 35*, eds S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh (Red Hook, NY: Curran Associates, Inc.), 27730–27744.
- Pang, B., and Lee, L. (2008). Opinion mining and sentiment analysis. *Found. Trends Inform. Retrieval* 2, 1–135. doi: 10.1561/1500000011
- Pennington, J., Socher, R., and Manning, C. D. (2014). “GLOVE: global vectors for word representation,” in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Stroudsburg, PA), 1532–1543. Available online at: <https://aclanthology.org/D14-1162/>
- Poquet, O., and Dawson, S. (2015). “Analysis of MOOC forum participation,” in *Asclite 15 - Conference of the Australasian Society for Computers in Learning in Tertiary Education* (Perth), 224. Available online at: <http://www.2015conference.ascilite.org/wp-content/uploads/2015/11/asclite-2015-proceedings.pdf>
- Poquet, O., Dowell, N., Brooks, C., and Dawson, S. (2018). “Are MOOC forums changing?” in *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (New York, NY), 340–349. Available online at: <https://dl.acm.org/doi/abs/10.1145/3170358.3170416>
- Prechelt, L. (2012). *Early Stopping — But When?*. Berlin; Heidelberg: Springer.
- Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., and Ré, C. (2017). “Snorkel: Rapid training data creation with weak supervision,” in *Proceedings of the VLDB Endowment. International Conference on Very Large Data Bases* (Rio de Janeiro: NIH Public Access), 269. Available online at: <https://dl.acm.org/doi/10.14778/3157794.3157797>; <https://dl.acm.org/toc/pvldb/2017/11/3>
- Ratner, A. J., De Sa, C. M., Wu, S., Selsam, D., and Ré, C. (2016). “Data programming: Creating large training sets, quickly,” in *Advances in Neural Information Processing Systems 29*, eds D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Red Hook, NY: Curran Associates, Inc.), 3567–3575.
- Ren, P., Xiao, Y., Chang, X., Huang, P.-Y., Li, Z., Gupta, B. B., et al. (2021). A survey of deep active learning. *ACM Comput. Surv.* 54, 1–40. doi: 10.1145/3472291
- Rosenthal, S., Farra, N., and Nakov, P. (2017). “SemEval-2017 task 4: sentiment analysis in Twitter,” in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (Stroudsburg, PA), 502–518. Available online at: <https://aclanthology.org/S17-2088/>
- Roy, A., Yee, M., Perdue, M., Stein, J., Bell, A., Carter, R., et al. (2022). “How COVID-19 affected computer science MOOC learner behavior and achievements: a demographic study,” in *Proceedings of the Ninth ACM Conference on Learning@ Scale, L@S’22* (New York, NY: Association for Computing Machinery), 345–349.
- Schmid, P. (2021). *Huggingface BERT-tweet-eval-emotion*. Available online at: <https://huggingface.co/philoschmid/BERT-tweet-eval-emotion>
- Schütze, H., Manning, C. D., and Raghavan, P. (2008). *Introduction to Information Retrieval*, Vol. 39. Cambridge: Cambridge University Press.
- Sennrich, R., Haddow, B., and Birch, A. (2015). Improving neural machine translation models with monolingual data. *arXiv preprint arXiv:1511.06709*. doi: 10.48550/arXiv.1511.06709
- Sha, L., Rakovic, M., Li, Y., Whitelock-Wainwright, A., Carroll, D., Gašević, D., et al. (2021). *Which Hammer Should I Use? A Systematic Evaluation of Approaches for Classifying Educational Forum Posts*. International Educational Data Mining Society.
- Sun, C., Li, S.-w., and Lin, L. (2016). “Thread structure prediction for MOOC discussion forum,” in *Social Computing: Second International Conference of Young Computer Scientists, Engineers and Educators, ICYCSEE 2016* (Harbin: Springer), 92–101.
- Sun, X., Guo, S., Gao, Y., Zhang, J., Xiao, X., and Feng, J. (2019). “Identification of urgent posts in MOOC discussion forums using an improved rCNN,” in *2019 IEEE World Conference on Engineering Education (EDUNINE)* (Lima: IEEE), 1–5. Available online at: <https://ieeexplore.ieee.org/abstract/document/8875845>
- Swinnerton, B., Hotchkiss, S., and Morris, N. P. (2017). Comments in MOOCs: who is doing the talking and does it help? *J. Comput. Assist. Learn.* 33, 51–64. doi: 10.1111/jcal.12165
- Topping, K. J. (2005). Trends in peer learning. *Educ. Psychol.* 25, 631–645. doi: 10.1080/01443410500345172
- Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., et al. (2023). Llama: open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*. doi: 10.48550/arXiv.2302.13971
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). “Attention is all you need,” in *Advances in Neural Information Processing Systems 30*, eds I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Red Hook, NY: Curran Associates, Inc.), 5998–6008.
- Wen, M., Yang, D., and Rose, C. (2014). “Sentiment analysis in MOOC discussion forums: What does it tell us?” in *Educational Data Mining 2014*, eds J. Stamper, Z. Pardos, M. Mavrikis, and B. M. McLaren (London: Citeseer), 130–137.
- Wise, A. F., and Cui, Y. (2018). Learning communities in the crowd: characteristics of content related interactions and social relationships in MOOC discussion forums. *Comput. Educ.* 122, 221–242. doi: 10.1016/j.compedu.2018.03.021
- Wise, A. F., Cui, Y., Jin, W., and Vytasek, J. (2017). Mining for gold: identifying content-related MOOC discussion threads across domains through linguistic modeling. *Internet Higher Educ.* 32, 11–28. doi: 10.1016/j.iheduc.2016.08.001
- Wong, J.-S., Pursell, B., Divinsky, A., and Jansen, B. J. (2015). “An analysis of MOOC discussion forum interactions from the most active users,” in *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction* (Heidelberg: Springer), 452–457. Available online at: https://link.springer.com/chapter/10.1007/978-3-319-16268-3_58#chapter-info
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y. (2020). A comprehensive survey on graph neural networks. *IEEE Trans. Neural Netw. Learn. Syst.* 32, 4–24. doi: 10.1109/TNNLS.2020.2978386

- Yamarik, S. (2007). Does cooperative learning improve student learning outcomes? *J. Econ. Educ.* 38, 259–277. doi: 10.3200/JECE.38.3.259-277
- Yang, D., Wen, M., Howley, I., Kraut, R., and Rose, C. (2015). “Exploring the effect of confusion in discussion forums of massive open online courses,” in *Proceedings of the Second (2015) ACM Conference on Learning@ Scale* (New York, NY), 121–130. Available online at: <https://dl.acm.org/doi/abs/10.1145/2724660.2724677>
- Yee, M., Roy, A., Stein, J., Perdue, M., Bell, A., Carter, R., et al. (2022). “The relationship between COVID-19 severity and computer science MOOC learner achievement: a preliminary analysis,” in *Proceedings of the Ninth ACM Conference on Learning @ Scale, L@S '22* (New York, NY: Association for Computing Machinery), 431–435.
- Zhang, L., Wang, S., and Liu, B. (2018). Deep learning for sentiment analysis: a survey. *Wiley Interdiscipl. Rev. Data Mining Knowledge Discov.* 8, e1253. doi: 10.1002/widm.1253
- Zou, W., Hu, X., Pan, Z., Li, C., Cai, Y., and Liu, M. (2021). Exploring the relationship between social presence and learners’ prestige in MOOC discussion forums using automated content analysis and social network analysis. *Comput. Hum. Behav.* 115, 106582. doi: 10.1016/j.chb.2020.106582



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Antonio Luque,
University of Almería, Spain
Najah Al-Shanableh,
Al al-Bayt University, Jordan

*CORRESPONDENCE

Julietta Noguez
✉ jnoguez@tec.mx

RECEIVED 22 June 2023

ACCEPTED 18 September 2023

PUBLISHED 02 October 2023

CITATION

Gonzalez-Nucamendi A, Noguez J, Neri L,
Robledo-Rella V and García-Castelán RMG
(2023) Predictive analytics study to determine
undergraduate students at risk of dropout.
Front. Educ. 8:1244686.
doi: 10.3389/feduc.2023.1244686

COPYRIGHT

© 2023 Gonzalez-Nucamendi, Noguez, Neri,
Robledo-Rella and García-Castelán. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Predictive analytics study to determine undergraduate students at risk of dropout

Andres Gonzalez-Nucamendi, Julieta Noguez*, Luis Neri,
Víctor Robledo-Rella and
Rosa María Guadalupe García-Castelán

School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico

In this work, a study is presented with quantitative variables using machine learning tools to detect undergraduate students at risk of dropping out and the factors associated with this behavior. Clustering algorithms and classification methods were tested to determine the predictive power of several variables regarding the dropout phenomenon on an unbalanced database of 14,495 undergraduate students with a real dropout rate of 8.5% and a retention rate of 91.5%. The usual classification criterion that assigns individuals to a class if their probability of belonging to it is greater than 50% provided accuracies of 13.2% in the dropout prediction and 99.4% in the retention prediction. Among eight classifiers, Random Forest was selected and applied along with Threshold Probability, which allowed us to gradually increase the dropout precision to more than 50%, while maintaining retention and global precisions above 70%. Through this study, it was found that the main variables associated with student dropouts were their academic performance during the early weeks of the first semester, their average grade in the previous academic levels, the previous mathematics score, and the entrance exam score. Other important variables were the number of class hours being taken, student age, funding status of scholarships, English level, and the number of dropped subjects in the early weeks. Given the trade-off between dropout and retention precisions, our results can guide educational institutions to focus on the most appropriate academic support strategies to help students at real risk of dropping out.

KEYWORDS

student dropout, learning analytics, predictive models, higher education, educational innovation, threshold probabilities, Artificial Intelligence

1. Introduction

The study of school dropouts is of interest at all educational levels. Reducing student dropouts is an important challenge that high schools and higher education must face. The loss of students who are beginning their high school or undergraduate studies constitutes a worldwide concern (e.g., Heublein, 2014; Aulck et al., 2016; Hsu and Yeh, 2019; Olaya et al., 2020). Several factors have been studied as the origins of dropping out, including unfavorable sociodemographic conditions, insufficient academic support, underprivileged economic income, and poor academic and social capabilities. Quantitative research on the

causes and the possible solutions for dropping out has been reported in the literature (e.g., [Aulck et al., 2016](#); [Garg et al., 2021](#)).

The difficulty in conducting extensive research on student dropouts is that many variables may play a simultaneously important role. For example, academic failure may be caused by a lack of clarity on the relevance of real-life scenarios ([Cameron and Heckman, 2001](#); [Wexler and Pyle, 2012](#)) and a lack of motivation that gives rise to random class attendance. Moreover, some students may have high rates of behavior problems because of a weak family and home structure ([Wexler and Pyle, 2012](#)). Some families place unrealistic expectations on their daughters and sons without providing them with the required tools for success. The absence of committed parents and family academic role models may also play a key factor ([Balfanz et al., 2007](#)). On the other hand, first-generation students may need a supportive environment in their schools to compensate for the sometimes-non-existent academic structure found at home.

To address this problematic situation, the Institute of Education and Science (US) has provided six recommendations to prevent dropouts at school ([Dynarski et al., 2008](#)):

- (1) Data systems should be aimed at the early detection of students at risk.
- (2) One-on-one tutoring is highly recommended for this population.
- (3) Academic support must be provided (extra office hours, extra homework, etc.).
- (4) Teaching social skills and providing specific programs to round out the class experience should not be underestimated.
- (5) Personalized academic instruction must be an option.
- (6) Focusing on lifelong competencies in addition to rigorous and relevant instruction must also be considered.

Studying the impact of the diverse factors that produce dropping out in middle and higher education has pushed institutions to perform statistical studies to disclose the relative importance of these factors and to apply suitable and timely measures to predict students at risk of dropping out ([Hsu and Yeh, 2019](#)). In this regard, the incorporation of learning analytics techniques that involve simultaneous analysis of students' social and performance data can disclose the factors that have a larger impact on dropping out. These techniques have contributed to the improved accuracy of predictive models in recent years (e.g., [Amare and Simonova, 2021](#); [Saravanan et al., 2022](#)). Nowadays, data analysis techniques are applied to large data sets to better understand the relationships among the multiple variables involved.

The present research should help improve the design of institutional retention programs by tailoring them to students who are at risk of dropping out but are the most likely to be retained according to appropriate selection algorithms. In this context, we use machine learning (ML) tools in terms of predictive analytics, to identify potential students at risk and define the characteristics that place them in such a situation. The information obtained will be useful in designing specific retention programs.

The objective of this paper is to find the most accurate predictive model that allows to make the best timely decisions for

institutional intervention, considering its ability to predict relative percentages of students at risk of dropping out.

The research questions that guide the present study are:

- (a) *What are the main factors that cause undergraduate dropout?*
- (b) *Which groups of students are the most vulnerable?*

In this first phase of this work, a study is presented with quantitative variables using machine learning tools to detect undergraduate students at risk of dropping out and the factors associated with this behavior. The organization of the paper is as follows. In section "2. Theoretical framework," a theoretical framework regarding the use of ML and learning analytics to predict dropping out is presented. In section "3. Related work," related studies in the literature on student dropouts are briefly summarized. Section "4. Methodology" presents the methodology followed in the present research and the case study selected. Section "5. Results and analysis" includes the principal results and analysis. Section "6. Discussion" presents the discussion, and finally, in section "7. Conclusion and future work," the conclusions and future work are outlined.

2. Theoretical framework

The machine learning (ML) tools and concepts used in this research are briefly described below.

2.1. Machine learning tools

2.1.1. Grouping or clustering algorithms

Clustering algorithms are procedures for grouping a series of vectors, associated with the variables according to specific criteria. Those criteria are usually distance or similarity. The closeness between the vectors is defined with a selected distance function, such as the Euclidean, although other metrics may be used. Generally, vectors in the same group (or clusters) share common properties. The knowledge of the groups allows a synthetic description of a complex multidimensional data set (e.g., [Romesburg, 2004](#)).

There are two main techniques for grouping: (a) hierarchical grouping, which can be agglomerative or divisive, and (b) non-hierarchical grouping, in which the number of groups is determined in advance, and the observations are assigned to the groups based on their closeness. For the latter technique, there are k -means and k -medoids methods.

2.1.2. The k -means method

The k -means method is probably the most used when the data set is so large that the computational time of the Hierarchical Clustering method, which is undoubtedly more accurate, is too large. In k -means, the number of groups is selected *a priori* and randomly creates an equal number of centroids; therefore, k -means does not always generate the same assignments for different program runs with similar conditions.

2.1.3. Predictive power of variables

To visualize and analyze the predictive power of a specific numerical variable and distinguish between dropping out and retention, this work applied the technique of density functions. It yields a continuous function derived from smoothing a histogram of relative frequencies, so the area under the curve represents probabilities. The diagrams in [Figure 1](#) illustrate this mechanism.

When it comes to a categorical predictive variable with k categories, the predictive power can be visualized by fusing a bar chart to distinguish among the categories with a greater or lesser proportion of dropouts, as shown in Section 3.1.5 below.

2.1.4. Classification methods

Classification methods are used to assign individuals to specific groups based on previously defined characteristics. In our study, the main characteristics associated with student dropout were (a) their academic performance during the early weeks of the first semester, (b) their average grade in the previous academic levels, (c) the previous mathematics score, and (d) the entrance exam score. Other important variables were the number of class hours being taken, student age, funding status of scholarships, English level, and the number of dropped subjects in the early weeks of the academic period. Algorithms determine the combination of these characteristics that define an individual's membership in a category. Predictive models are machine learning techniques applied to databases that seek to identify patterns to predict the membership of individuals in categories and make informed decisions. The predictive area has recently assumed a leading role in education (e.g., [Liu et al., 2022](#)).

We have selected eight classifiers from a wide range of available options, based on our previous experience and the diversity of approaches they offer. These classifiers were chosen specifically to address our classification problem. The list includes Support Vector Machine (SVM), which searches for a separating hyperplane in a feature space ([Cortes and Vapnik, 1995](#)); K-Nearest Neighbors (KNN), which classifies based on closeness to the nearest K data points ([Cover and Hart, 1967](#)); Decision Trees, a hierarchical structure that makes classification or regression decisions using nodes representing feature questions ([Quinlan, 1986](#)); Random Forest, an ensemble of decision trees that combines results to improve accuracy ([Breiman, 2001](#)); ADA Boosting (Adaptive Boosting), an ensemble that improves weak classifiers by assigning greater weight to incorrectly classified instances ([Freund and Schapire, 1996](#)); Extreme Gradient Boosting (XGBoost), an efficient implementation of boosting with multiple decision trees ([Chen and Guestrin, 2016](#)); Naive Bayes, a probabilistic classifier based on Bayes' theorem that assumes independence among features ([Duda et al., 2001](#)); and LDA (Linear Discriminant Analysis), which finds linear combinations of features for discrimination between classes ([Fisher, 1936](#)). These classifiers were selected for their versatility and ability to address a wide variety of approaches to solving our problem.

The evaluation of a predictive model is based on a confusion matrix, which is a valuable tool to assess how well an ML classification model works. It is used to show explicitly when one class is confused with another, which allows working separately with different error measures. *Positive precision* refers to the dropout cases and *negative precision* refers to the retention cases.

Therefore, these values and the *overall accuracy* of the prediction can be obtained as follows:

Positive precision: Percentage correctly classified as dropout;

$$PP = TP / (TP + FP).$$

Negative precision: Percentage correctly classified as retention;

$$NP = TN / (FN + TN).$$

Overall accuracy: Percentage the total number of cases correctly classified:

$$OA = (TP + TN) / (TP + TN + FP + FN).$$

Although F-Measure is commonly used to compare classifiers ([Powers, 2020](#)) it has limitations in situations of unbalanced classes and varying probability thresholds. This is because it varies by threshold, which makes comparisons difficult. In addition, it is sensitive to class imbalance and may bias the evaluation toward the majority class. For these reasons, we chose to visually assess the performance of the eight classifiers through a scatter plot ([Figure 2](#)) showing the probability of dropout on the X -axis and the probability of retention on the Y -axis. We observe that Random Forest, with high "accuracy," stands out as a promising option that requires less tuning to achieve good performance. This leads us to prefer Random Forest over other classifiers that demand exhaustive hyperparameter settings.

Additionally, Random Forest is a technique of great importance in the analysis of dropout data since it allows us to visualize the importance of the predictor variables. We can obtain a graph showing the relative importance of the variables and their individual effect on model improvement, i.e., how much the overall accuracy of the model is damaged by considering the absence of each variable in the whole forest. The most important variable is assigned 100%, and the others are given relative importance in the form of a number between 0 and 1. Then, a standardization is performed so that the sum of all contributions equals 100%.

2.1.5. Threshold probability method as assignment criterium

When the variable to be predicted is highly unbalanced, as is the case for the retention and dropout cases in the example shown in [Figure 3](#), a bias toward the dominant class may occur, even when classifying all individuals in that class. This usually happens when using a fixed probability threshold. For example, we can consider that a variable to predict A , has two categories: Yes or No. Traditionally, $A = \text{Yes}$ is assigned whenever the probability $P(A = \text{Yes}) > 0.5$; and $A = \text{No}$, is assigned otherwise. This logic works well when the training data is balanced, that is, when it contains approximately the same number of Yes and No cases. However, this does not happen in real scenarios where there is a large imbalance. For example, suppose that a database has only 10% of Yes cases and 90% of No cases. Under these circumstances, the overall precision measure is misleading because a naïve (and useless) rule assigning all cases to No would have a global precision of 90%, with a precision of No at 100% but a precision of Yes at 0%. Generally, the accuracy of Yes is the most interesting in real cases and therefore, in the given example, this assignment would be useless.

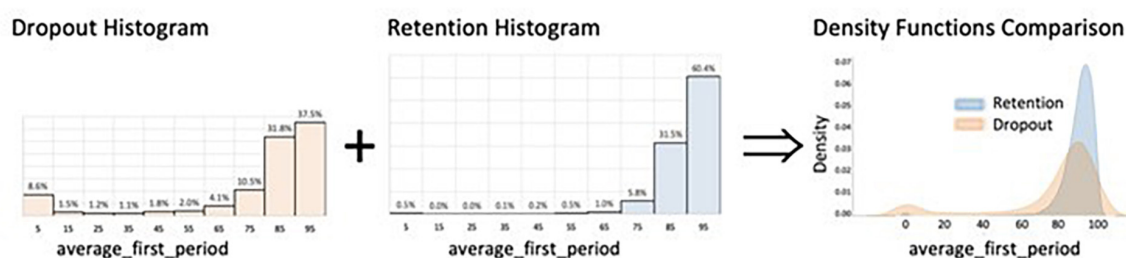


FIGURE 1
Example illustrating density functions.

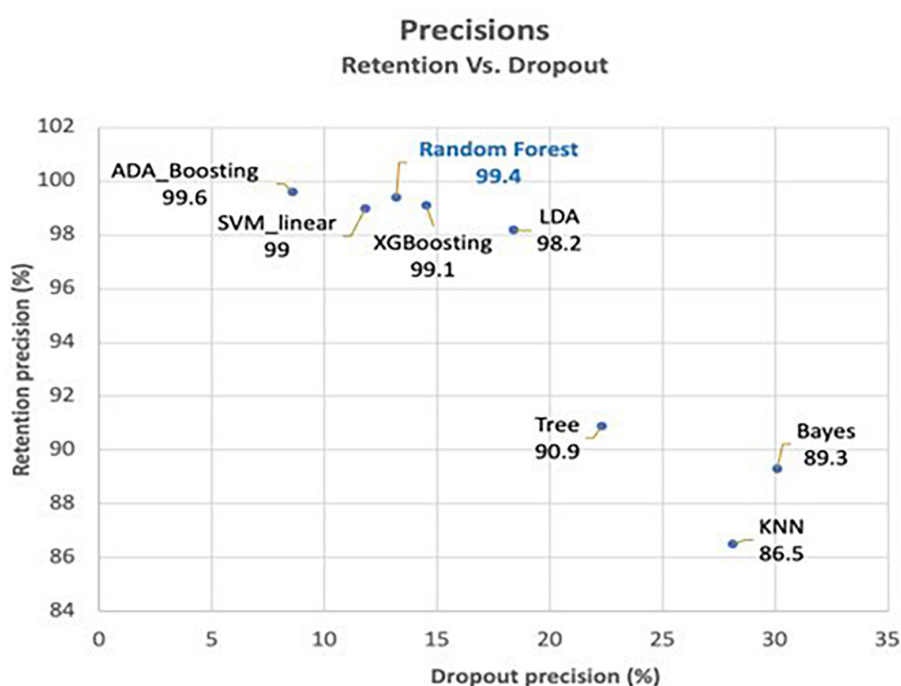


FIGURE 2
Comparison of retention and dropout precision classification methods. The retention percentage is indicated.

One way to reduce this problem is by varying the *cut-off probabilities*, requiring more probability from the class with the larger number of cases, and consequently, requiring less probability from the class with fewer cases. For example, we could manually set the rule $P(A = \text{Yes}) \geq 0.2$ to assign Yes and consequently $P(A = \text{No}) > 0.8$ to assign No, and a better balance for the prediction probabilities for both classes would be achieved. Nevertheless, it is not recommended to assign such a low threshold chance to the non-dominant class to attain an accuracy closer to 100%, since it would be at the cost of huge damage to the dominant class accuracy (see Section “5.3.3. Predictive power of explicative variables” below).

2.2. Dropouts

Dropping out is a situation in which the student withdraws from an educational institution or system without obtaining

accreditation or a school certificate (e.g., [Lamb et al., 2010](#)). It can occur at any educational level and is a complex problem due to many endogenous and exogenous variables, as presented in section “4. Methodology.” Endogenous variables refer to the intrinsic characteristics of students, for example, their ability to learn, their interest in school, or their level of development. Exogenous variables are related to external factors such as economic factors, family conditions, and natural disasters.

3. Related work

Dropout models deal with complex issues in which individual choices, institutional processes, demographic background, health issues, teachers’ opinions, student behavior and social factors play a role when a student decide to whether or not remain at the University ([Hedge and Prageeth, 2018](#)). The inability to cope with the performance demands of higher education institutions, wrong

expectations, financial problems, keeping pace with lecturers, rowdy classrooms, time management and less identification with the career path are the most important reasons for dropping out (Aulck et al., 2016; Govender, 2020). GPAs in math, English, chemistry, and psychology, as well as birth year were among the strongest predictors of student persistence (Aulck et al., 2016). Student records and transcripts for courses taught in the first 2 years, high school averages, and whether the student graduated from the chosen major or not are all valuable input variables in the dropping out understanding (Abu-Oda and El-Halees, 2015; Von Hippel and Hofflinger, 2020). High school performance in humanities has a surprisingly significant impact even on engineering students (Nagy and Molontay, 2018).

Considering all the mentioned issues, Germany has considered academic policies that include broad assistance measures, such as flexibilization of the curricula, better information for students, and the expansion of support offered at the start of their studies (Heublein, 2014). Palestinian studies found out that digital design and algorithm analysis have a great effect on predicting student persistence in the major and decreasing the likelihood of students dropping out (Abu-Oda and El-Halees, 2015).

Some other universities around the world have used Machine Learning techniques, Naive-Bayes Classification Algorithms programmed in R, Gradient Boosted Trees, Deep Learning, rough set theory and k-means in an effort to determine the factors that influence dropping out (e.g., Abu-Oda and El-Halees, 2015; Aulck et al., 2016; Hedge and Prageeth, 2018; Nagy and Molontay, 2018; Olaya et al., 2020; Von Hippel and Hofflinger, 2020). Even Thematic analysis has been used to analyze qualitative narrative data (e.g., Govender, 2020).

Table 1 shows a comparison of the related work outlined above. In column 2 the statistical technique used, or the approach followed, by the different authors are outlined. In columns 3 to 10 the most relevant dropout factors extracted from among these references are indicated. The “x” signs indicate the dropout factors considered in each reference. From this table it is seen that the three most common and relevant factors for dropping out are: (a) first-year undergraduate grades, (b) high school grades, and (c) university entrance exam scores.

The following section describes the research methodology used in this study.

4. Methodology

To build predictive models to identify high-risk students in a timely manner we followed the research methodology indicated in **Figure 4**: (a) Case study selection; (b) Data cleansing and definition of the study variables; (c) Identification of relevant database subsets; (d) Definition of the research hypothesis; (e) Application of statistical and ML techniques; (f) Results and analysis; (g) Discussion; and (h) Conclusions.

4.1. Case study

As a case study, the analysis of dropout cases between 2014 and 2021 for a prominent private university in Mexico was chosen.

Approximately 8.1% of the students who entered this institution did not manage to finish their studies or transferred to other institutions (Alvarado-Uribe et al., 2022). Although this dropout rate is low compared to the average for other Mexican universities, it does represent an important social cost and economic effort for families to support their sons' and daughters' studies. Moreover, school fees may be absorbed not only by parents or families but also by other institutions that regularly provide scholarships. In the second phase of this work, we are expanding the study to include qualitative variables such as socio-economic categories and social lag, that will be reported in a future work.

4.2. Data cleansing

We analyzed an institutional initial database of 143,326 records (students) with 50 independent variables (Alvarado-Uribe et al., 2022). To proceed with this research, a careful study and cleansing of the initial database yielded a suitable database for applying the selected ML techniques.

4.3. Sample breakdown

The initial database contained data from 2014 to 2022 including high school and college students. The institution launched a new educational model at the undergraduate level (Tec21, 2022) in the fall of 2019 (August–December 2019), so this research focuses on first-year undergraduate students enrolled in this new educational model to determine the variables that most influence dropouts and to propose intervention schemes.

4.4. Hypothesis

Derived from the research questions, the following hypothesis was established:

1. *It is possible to identify in a timely manner the key differentiating characteristics of undergraduate dropouts, and to cluster students for timely and adequate support.*

4.5. Machine learning analysis strategies

The ML analysis strategies considered were: (a) clustering; (b) classification methods comparing populations of dropouts and non-dropouts, where eight classification techniques were considered; (c) Random Forest in detail and Threshold Probability Method (TPM); and (d) the predictive power of the variables. The following section shows and analyzes the results.

5. Results and analysis

According to the established methodology (**Figure 4**), the most important results of each step are described below.

TABLE 1 Comparison of related work.

References	Technique	University first-year academic records	Entrance exam scores	High School grades	First choice major denied	Unfavorable sociodemographic conditions	Wrong expectations	Underprivileged economic income	Insufficient academic support
Relevant dropout factors									
Heublein, 2014	Empirical research	×	×	×	×	×	×	×	×
Abu-Oda and El-Halees, 2015	Decision Tree, Naive-Bayes, <i>k</i> -means, linear models, deep learning	×	×	×	×				
Aulck et al., 2016	Regularized logistic regression, <i>k</i> nearest neighbors, random forest	×				×	×	×	×
Hedge and Prageeth, 2018	Decision tree, Naive-Bayes, <i>k</i> -means, linear models, deep learning	×	×	×	×				
Nagy and Molontay, 2018	Decision tree-based algorithms, Naive Bayes, <i>k</i> -NN, linear models, and deep learning			×					
Hsu and Yeh, 2019	Hybrid approach: <i>k</i> -means, set theory	×	×	×					
Olaya et al., 2020	Uplift modeling	×	×	×	×	×	×	×	×
Von Hippel and Hofflinger, 2020	Simple logistic regression	×	×	×	×				
Govender, 2020	Thematic analysis	×					×		×

TABLE 2 Cluster characteristics.

Cluster number	Cluster size (students)	Retention (students)	Dropouts (%)
Cluster 0	8,230	7,713	6.3
Cluster 1	743	621	16.4
Cluster 2	5,522	4,927	10.8

5.1. Database description

As mentioned above, a curated database (DB) provided by our Institution was used as the data source (Alvarado-Uribe et al., 2022). This initial database included $N_{tot} = 143,326$ students and 50 academic/demographic variables. The DB contained information on 65,809 high school students and 77,517 undergraduate students enrolled from August–December 2014 to August–December 2020.

Supplementary Appendix Table 1 presents the 16 variables selected in the first phase of the study from the initial database. The variable name, their description, and their type are specified. Numerical variables use a continuous range of values within a given numerical interval, while categorical variables use a discrete set of data. The 16 variables used in this study include 14 numerical and 2 categorical.

5.2. Undergraduate student sample

The following analysis is divided into two parts: (a) a description of the undergraduate students' sample and the cleansing process, and (b) the derived results applying different ML analysis strategies to the cleaned sample. Both the data cleansing and the algorithm execution were programmed in Python using NumPy, Pandas, Matplotlib, and Scikit-learn libraries.

5.2.1. Cleansed used variables

Although the original sample consisted of 143,326 students, this research focused only on the 77,517 students in the undergraduate sample. From the undergraduate subset, only the 24,507 first-year students enrolled in the educational model (Tec21, 2022) at the Institution were considered. However, when making the selection of the 16 numerical variables, it was identified that many students did not have defined values for these variables, so it was necessary to eliminate those students from the sample. The homogeneous sample without empty entries considered 14,495 complete records. This is the final cleansed sample to which the machine learning analysis strategies described below were applied. It is important to state that the variable to be predicted in this research is the *retention* variable (number 16 in Supplementary Appendix Table 1).

5.3. Machine learning analysis strategies

The ML analysis strategies comparing dropout and non-dropout populations are: (a) clustering, (b) classification methods, (c) Random Forest in detail with Threshold Probability Method (TPM), which is helpful for unbalanced data classification (Rodríguez Rojas, 2022), and (d) predictive power of the variables.

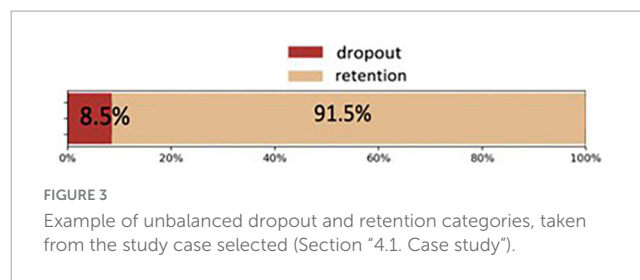


FIGURE 3

Example of unbalanced dropout and retention categories, taken from the study case selected (Section "4.1. Case study").

In (b), eight classification techniques were considered: (1) Support Vector Machine (SVM), (2) k -Nearest Neighbor (KNN), (3) Decision Trees (DT), (4) Random Forest (RF), (5) Adaptive Boosting (ADA_Boosting), (6) Extreme Gradient (XG_Boosting), (7) Bayesian Classifier (BC), and (8) Linear Discriminant Analysis (LDA). Below are the main results.

5.3.1. k -means clustering

To explore the prediction through cluster formation, the cleansed database was segmented into three main clusters, using the k -means technique. Cluster 0 had 8,230 students, 7,713 were retained and 517 dropped out. Cluster 1 had 743 students, 621 were retained and 122 dropped out. Cluster 2 had 5,522 students, 4,927 were retained and 595 dropped out. The corresponding sizes, retention numbers, and dropout percentages are shown in Table 2. The characteristics of the clusters are analyzed below.

In Figure 5, the cluster sizes are presented in a pie chart, and in Figure 6 the corresponding dropout percentages are shown in a bar chart. In Figure 7, a radar diagram emphasizes the main variables related to student dropouts for each cluster. The radar diagram is normalized to amplify the differences among clusters, giving values between 0 and 100% to those variables with the lowest and largest weights, respectively. Figure 7 presents the 14 explicative variables and the *dropout_semester* variable.

It should be noted that the information provided by the *dropout_semester* variable is equivalent to that of the retention variable. In fact, the value *dropout_semester* = 0 is equivalent to retention, and values *dropout_semester* = 1, 2, 3, or 4 are equivalent to dropout, that is, the *dropout_semester* variable is the breakdown in semesters of the dichotomous variable to be predicted, retention. Therefore, Figure 7 shows the relative weight among clusters of the 14 explicative variables selected in this research to explain the variable to be predicted (*dropout_semester* or *retention*).

Figure 6 shows that Cluster 0, representing 38% of the student sample, has the lowest dropout percentage, 6.3%. This cluster is characterized by students who (see Figure 7) have: (a) an intermediate percentage of dropped subjects or failed subjects during the first period of the first semester, (b) the highest average grade in the first period of the first semester, (c) the highest percentages of scholarship and loans, (d) the highest percentage of full-time students, (e) the highest general math evaluation, admission rubric score, and English evaluation, (f) the lowest percentage of students who took the admission test online, (g) the highest admission test and previous-level average-scores, and (h) the youngest students of the sample.

Cluster 1 is the smallest (5% of the student sample) and has the highest dropout percentage (16.4%). This cluster is characterized by: (a) the oldest students in their class, (b) the highest percentage

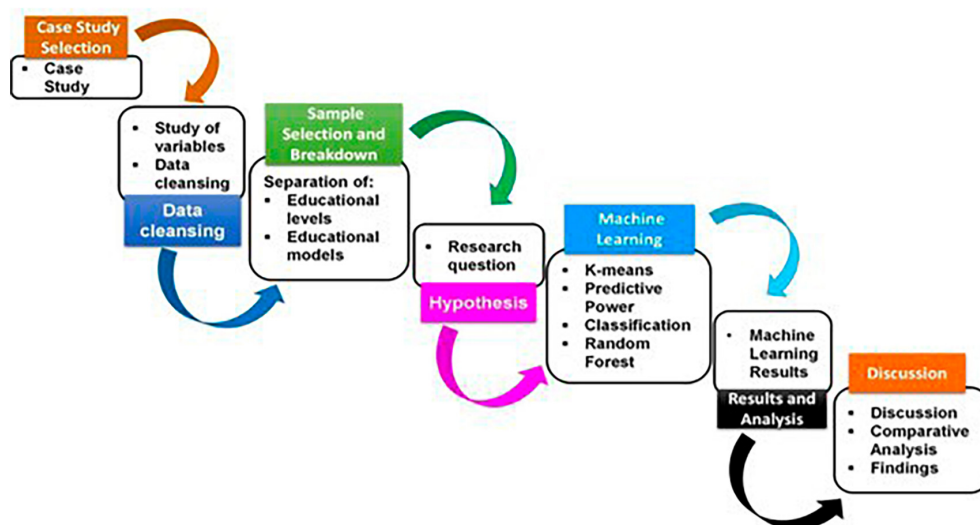


FIGURE 4
Steps of the research.

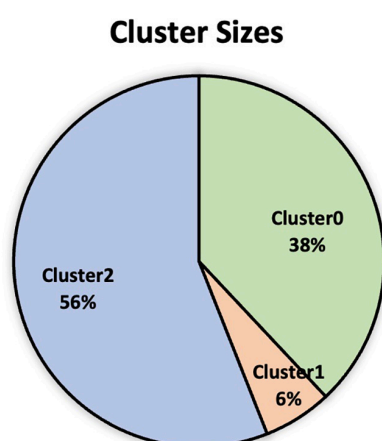


FIGURE 5
The cluster sizes.

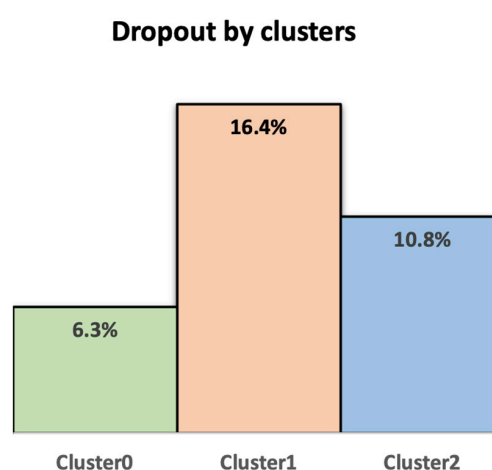


FIGURE 6
Dropout percentages by cluster.

of dropouts during the first year (*dropout_semester*), (c) relatively low percentages of dropped subjects or failed subjects during the first period of the first semester, (d) a low average grade in the first period of the first semester, (e) low percentage of scholarship or loans, (f) low percentage of full-time students, (g) a very low general math evaluation, admission rubric score, and English evaluation, (h) the highest percentage of students taking the admission test online, and (i) the lowest admission test and lowest previous level average scores.

Finally, Cluster 2 represents 57% of the student sample and has a relatively high 10.8% dropout percentage. This cluster is characterized by students: (a) with a high percentage of dropped subjects and failed subjects in the first period of the first semester, (b) the lowest average grade during the first period of the first semester, (c) no significant scholarships or student loans, (d) an intermediate percentage of full-time students, (e) general math evaluation, admission rubric and English evaluation that were

average within the student sample, (f) a low percentage taking the admission test online, (g) admission test and previous-level average-score at an intermediate level, and (h) an age between those of clusters 0 and 1.

5.3.2. Classification methods

Several ML classifiers were tested to obtain the best accuracy for dropping out, retention, and/or global percentages. Through the execution of by-default parameters that required just a few adjustments, the classifiers gave dropout, retention, and global precision percentages for the undergraduate students in the sample ($N = 14,495$). The results are shown in Table 3. The comparison of the precisions obtained for the retention and dropout percentages by each classifier is shown in Figure 2 above.

The graph shows a negative relationship between the percentages of dropout precision and the percentage of retention

TABLE 3 Comparison of classifiers with by-default parameters.

Classifiers	Precision percentage (%)		
	Global	Dropout	Retention
SVM_linear	91	11.8	99.0
KNN	82	28.1	86.5
Decision Tree	85	22.3	90.9
Random Forest	92	13.2	99.4
ADA_Boosting	92	8.6	99.6
XGBoosting	92	14.5	99.1
Bayes	84	30.1	89.3
LDA	91	18.4	98.2

TABLE 4 Random Forest Dropout, Retention, and Global prediction precisions for different threshold probabilities.

Threshold probability	Dropouts	Retention	Global
0.000	0.000	1.000	0.915
0.400	0.060	0.997	0.917
0.500	0.120	0.995	0.920
0.600	0.181	0.990	0.920
0.700	0.233	0.983	0.918
0.750	0.257	0.970	0.909
0.880	0.510	0.811	0.785
0.889	0.534	0.783	0.761
0.901	0.586	0.745	0.731
0.945	0.755	0.480	0.504
0.995	1.000	0.019	0.103
1.000	1.000	0.000	0.086

precision. Overall, increasing the precision in dropout percentage yields a decreasing precision in the retention percentage. This fact can be used to obtain the optimum combination according to the requirements of each educational institution. There are only two exceptions to this rule. The first is in which RF (Random Forest) produces both higher dropout and retention precisions than SVM. The second is found in the lowest part of the retention precisions, where Bayes dominates over KNN, with higher precisions both in dropping out and retention. Notice that in the graph, all percentages of retention precision are greater than those of dropout precision. This is due to the imbalance in the variable that will be predicted (dropout percentage vs. retention percentage: 8.5% vs. 91.5%, respectively) as can be seen in [Figure 3](#) above.

5.3.3. Random forest and threshold probability method (TPM)

Using the Random Forest (RF) classifier with $n_estimators = 400$, $criterion = gini$, $min_samples_split = 18$, and changing the threshold probabilities, a threshold-probability graph was obtained. The results are shown in [Table 4](#). [Figure 8](#) plots the graphs of the dropout precision, retention precision, and

TABLE 5 Confusion matrix used to calculate effectiveness coefficients.

		Dropout	Retention	Real totals
Prediction				
Actual values	Dropout	127	122	249
	Retention	501	2,149	2,650
	Predicted totals	628	2,271	2,899

global precision of the sample. Threshold probability is along the horizontal axis and the vertical axis corresponds to the precisions (dropout, retention, and global). The training sample contained 80% of the cases and the remaining 20% corresponded to the testing sample. To obtain meaningful values for the comparison, the random seed was fixed as $random_state = 0$.

For an educational institution, it is possible to take advantage of these results to plan how to distribute the resources in retention efforts. For example, if the intervention point is selected at 51% dropout precision and 81% for retention according to [Table 4](#) (or [Figure 8](#)) the threshold probability is 0.88, and the global precision for the total sample is 78.5%. A measure that can be obtained with these results is the *effectiveness coefficient*, defined as the expected number of effective interventions that the institution should offer to students correctly predicted as dropouts divided by the total number of interventions the institution would offer to any student predicted as dropout (correctly or incorrectly) according to the model. To explain this coefficient, we use the testing sample that results from the remaining 20% of the students that were not included in the training sample. The total number of records in the testing table is therefore $0.2 \times 14,495 = 2,899$. The corresponding confusion matrix is presented in [Table 5](#).

Out of the 249 real dropouts in the database for the testing sample, 127 were correctly detected and addressed, but 122 were undetected and, consequently, left unattended. Therefore, the precision in dropouts is $127/249 \approx 51\%$. Similarly, out of the 628 students predicted as dropouts, only 127 were true dropouts and 501 were false dropouts. The expected effectiveness is then $127/628 \approx 20\%$. If academic institutions implement intervention programs to attend to this population at risk of dropping out, only 1 out of 5 students will need these programs, while the remaining 4 will not, leading to misspending of valuable academic and economical resources. Nevertheless, note that for a random intervention the effectiveness would still decrease to only 8.5%, which is the total dropout percentage for the entire $N = 14,495$ undergraduate student sample, representing an even greater waste of academic and economic resources. Therefore, due to the imbalance among the types of variables to be predicted, it is useful to apply the threshold-probability method, because it can vary the precisions in the prediction of the dropout and of the retention class. This can guide institutions to implement the best interventions to address dropout cases.

5.3.4. Random forest important variables

Different tests were applied with several classifiers. Random Forest (RF) was chosen because of its good performance in retention and dropout percentages precisions, and because it

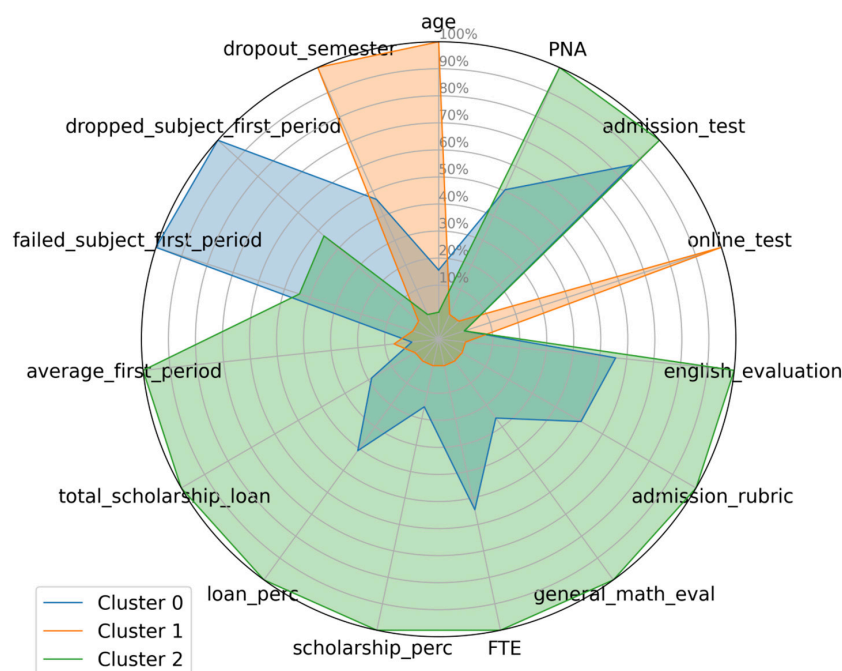


FIGURE 7

Cluster comparison for the selected variables (see [Supplementary Appendix Table 1](#) for variable names).

provided information on the importance of the used variables. [Figure 9](#) shows a sketch with the relative importance of the variables used, according to RF. It is important to notice here that only 11 of the 14 explicative variables were selected in RF, excluding those that complicated or even damaged the precision of the classifier. It is seen that, according to the RF classifier, the most important variable associated with student dropout is the average grade obtained in the first period of the first semester. Other important variables are: (a) the previous level average score (PNA), (b) the results of the general math evaluation (of the admission test and/or from the school of origin), (c) the admission test and admission rubric results, (d) full-time student status (FTE), (e) the student's age (younger students have lower dropout percentages than older ones, as mentioned in section “5.3. Machine learning analysis strategies”), (f) the total scholarships and student loans, and (g) the English evaluation result. According to RF, less important variables are the number of dropped out subjects in their first period or if the student took the admission test online.

5.3.5. Predictive power of explicative variables

To better know the influence of each variable in the dropout prediction, density function comparisons were made for the dropout class and the retention class, as explained in section “2.1.3. Predictive power of variables” (see [Figure 1](#)). The diagrams are presented in [Figure 10](#), in order of importance according to the RF classifier. These diagrams reinforce the results already presented in [Figure 9](#). The horizontal axis represents the range of possible values of the variables and the vertical axis shows the relative importance of that variable for predicting retention (in blue) and dropout (in peach color) cases. For example, from [Figure 10.1](#), if students obtain in their first period a grade higher than 90, they will most probably be retained. On the other hand, if their grade was lower than 80, they were more likely to drop out. Similarly, [Figure 10.2](#)

Precisions vs. Threshold Probability

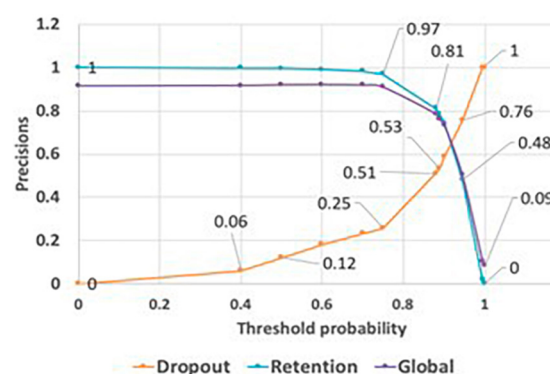


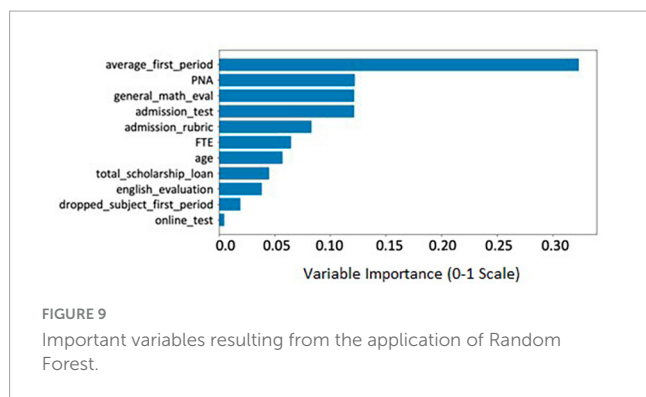
FIGURE 8

Random Forest Dropout, Retention and Global precisions vs. threshold probability.

suggests that students with previous-level average grades (PNA) higher than 90 will likely be retained, while students with previous-level average grades below 80 are more likely to drop out. Similar conclusions can be seen in [Figures 10.3–10.12](#). Bumps in the x-axis may correspond to input variables discretization.

6. Discussion

There are several methods to address the problem of class imbalance in the context of machine learning ([Douzas et al., 2018](#)), which refers to the situation where one of the classes



in a dataset is significantly smaller than the other. Many of these methods involve generating synthetic data, either through under sampling (removing records from the majority class) or oversampling (creating records in the minority class). However, this strategy raises the concern that it may distort the reality of the data, which in turn could affect the accuracy of algorithms when faced with new data.

An important observation is that, to date, we have not found references in the literature that use the probability threshold method in the context of student dropout. This suggests a scarcity in the application of this method in this particular area. Therefore, one of the main contributions of this article lies in presenting the probability threshold approach to address the problem of predicting student dropout.

The probability threshold approach involves varying the cutoff probability in the assignment criterion to either class. In the context of student dropout, this method offers an additional advantage, as it allows for the modulation (or selection) of the balance between “Yes” rates (Recall) and “No” rates (Sensitivity) according to convenience and the associated costs for the educational institution.

The results obtained from the ML techniques applied to the selected sample of 14,495 undergraduate students consistently showed that the average grades in the first university period (5 weeks), the admission tests, and the average grades in high school are the three most important variables to predict undergraduate dropouts. This is in line with research reported by different authors as presented in the literature review (Table 1). For instance, Abu-Oda and El-Halees (2015), Hedge and Prageeth (2018), Olaya et al. (2020), and Von Hippel and Hofflinger (2020) reported that variables such as *entrance exam scores*, *high-school grades*, *University*, and *first-year academic records* are relevant for dropout prediction. Our findings are based on the study of the importance of variables obtained with the Random Forest classifier, clustering analysis, and the study of predictive power through density functions. The results from this work strongly suggest detecting students with high-risk dropout timely in the first weeks of the first academic term. Organizing additional individualized tutoring or workshops to support students with high-risk dropout characteristics during this period should be implemented as soon as possible.

Of the eight classifiers explored in this research, Random Forest (RF) provided the highest percentages of accuracy for the total sample of students, the students who dropped out, as well as for the

students who were retained. The results found in this study indicate that, given the imbalance between the dropout and retention variable percentages in our student sample (8.5% and 91.5%, respectively), the best threshold value is not the one that gives the best accuracy for the whole sample to predict retentions and dropouts, but the one that gives the best precision in determining dropouts *while still* maintaining an acceptable precision in the retention and global precisions (Figure 8). The equilibrium point (where the three curves intersect) shows that it is possible to attain dropout precision close to 0.70 while also maintaining the retention and the global precisions at about the same value. Institutions may consider this value to better determine high-risk students and implement more focused actions to attend to this population, making the implemented resources more efficient. On the contrary, if global precision as high as 0.92 is adopted, the dropout precision would be only about 0.12 (Table 4), missing a great majority of high-risk students. While it is always possible to find rules that can classify any individual class with 100% precision, this comes at the expense of losing precision for another one, as shown in the Precision vs. Threshold probability graph (Figure 8).

It is worth noting that the database employed for this study contains dropout information only from the first academic year. Therefore, a limitation of this research is the lack of data for students dropping out in the third or higher semesters of their academic programs, although it is likely that the corresponding numbers are lower than for the first two terms because these students would be expected to be more adapted to the characteristics of their programs. More information about students after their second year is required to fully understand the long-term effect of the intervention efforts from academic, social, and economic perspectives.

Currently, we are expanding the study to include qualitative or categorical variables such as socioeconomic variables, honors and award, scholarships, social lag, parental education, etc. This second part of the research will be published elsewhere (see Gonzalez-Nucamendi et al., 2023).

The design of specific retention programs is beyond the scope of this study, and they are expected to include a comparative study of experimental (with academic intervention programs) and control groups over the years. These programs should prove the virtues of the models of improvement in tailoring retention efforts in middle and higher education over conventional predictive modeling approaches (e.g., Olaya et al., 2020).

7. Conclusion and future work

Through various Machine Learning techniques, the main variables associated with first-year undergraduate student dropouts in 14,495-student sample of the selected Case Study were identified. The most relevant numerical classification variables were: (a) the student's academic performance in the first weeks of the first semester, (b) the average grades of the previous academic level, (c) the general entrance score in mathematics, and (d) admission test results. Other important variables included: (a) the number of class hours, (b) the age of the student, (c) the scholarship, (d) the English level, and (f) the number of subjects dropped in the first weeks of the term.

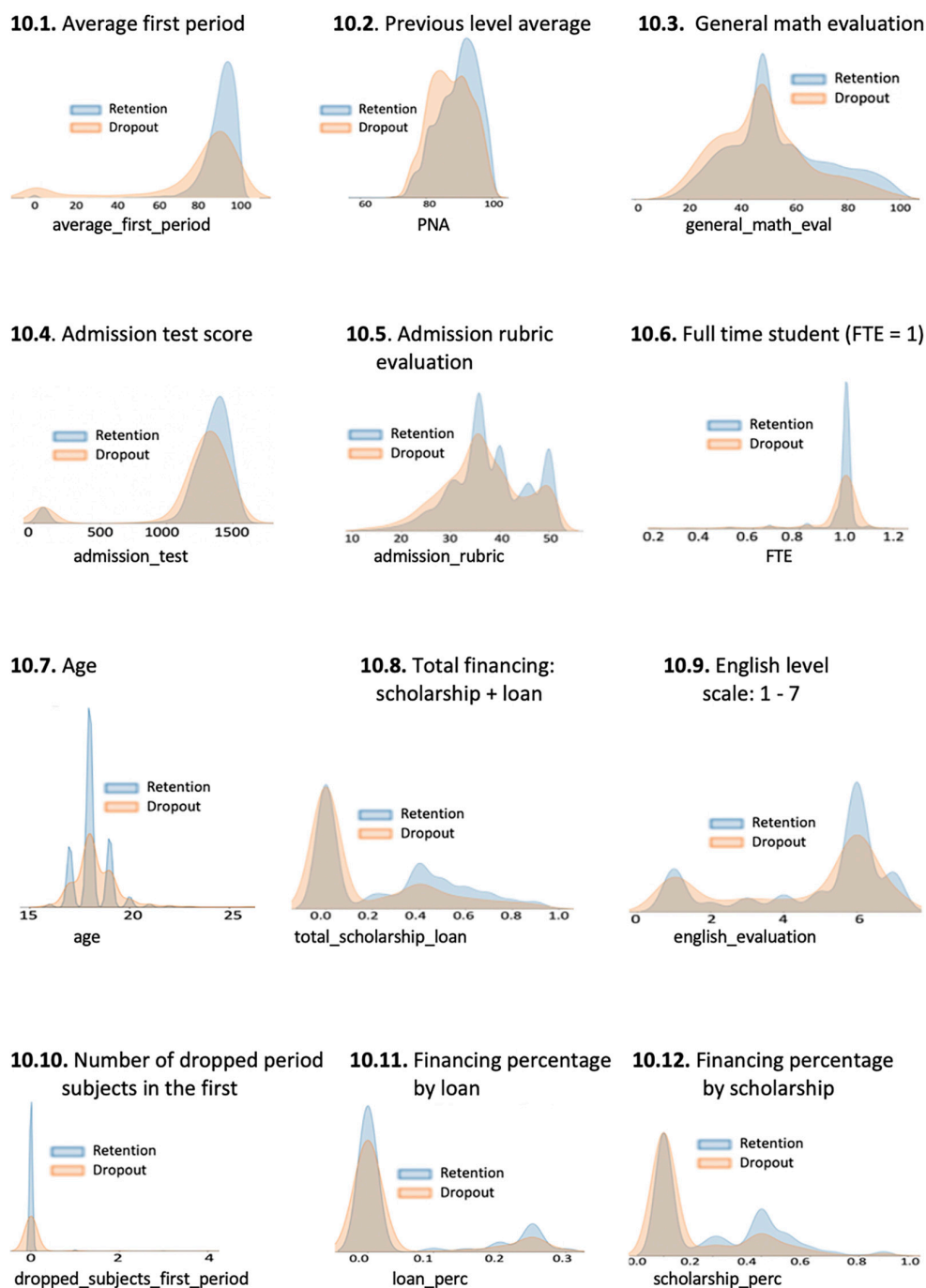


FIGURE 10
Density functions comparison for dropout and retention predictive variables.

Among the eight classifiers explored in the analysis of the Case Study data of this research, the Random Forest (RF) classifier provided the highest percentages of accuracy for the total sample of students, the students who dropped out, as well as for the students who were retained. Analyzing the predictions obtained with various classification algorithms, a negative relationship was found between the accuracies in predicting dropout and retention percentages. This led us to the use of a probability threshold different from 50% as a classification criterion to favor the smallest class and achieve a better balance in the prediction accuracy between

unbalanced classes. This resulted in an improved accuracy in detecting dropouts. With this, a control is also provided that allows regulating the dropout and retention precision levels to achieve flexibility so that universities can adapt them to their objectives, resources and needs. In the database analyzed, the use of the Random Forest algorithm to implement the Threshold Probability methodology resulted in the most appropriate approach.

Consequently, the results for the Case Study of this research clearly show that the best strategy is not the one that provides the best overall prediction accuracy for the whole student sample,

but the one that predicts the highest accuracy in dropout percentage while still maintaining appropriate overall and retention probabilities precision.

The design and the implementation of segmented or personalized interventions are better than random, non-focalized interventions. In this sense, academic institutions should provide appropriate programs to offer tutoring and support primarily to those students early detected as possible dropout candidates, to increase their retention probabilities.

Data availability statement

The data that support the findings of this study are available from the Institute for the Future of Education (IFE)'s Educational Innovation collection of the Tecnológico de Monterrey's Data Hub, but restrictions apply to the availability of these data, which were used under a signed Terms of Use document for the current study, and so are not publicly available. Data are however available from the IFE Data Hub upon reasonable request at <https://doi.org/10.57687/FK2/PWJRSJ> (accessed January 28, 2023).

Ethics statement

Ethical approval was not required for the study involving human data in accordance with the local legislation and institutional requirements. Written informed consent was not required in accordance with the national legislation and the institutional requirements.

Author contributions

AG-N, JN, LN, VR-R, and RG-C contributed to the conception and design of the study and wrote the first draft of the manuscript. VR-R organized the database. AG-N performed the statistical analysis. All authors contributed to the manuscript revision, read, and approved the submitted version.

References

- Abu-Oda, G. S., and El-Halees, A. M. (2015). Data mining in higher education: University student dropout case study. *Int. J. Data Mini. Knowl. Manag. Process* 5, 15–27. doi: 10.5121/ijdkp.2015.5102
- Alvarado-Urbe, J., Mejía-Almada, P., Masetto Herrera, A. L., Molontay, R., Hilliger, I., Hedge, V., et al. (2022). Student dataset from tecnológico de monterrey in Mexico to predict dropout in higher education. *Data* 7:119. doi: 10.3390/data7090119
- Amare, M. Y., and Simonova, S. (2021). "Global challenges of students' dropout: A prediction model development using machine learning algorithms on higher education datasets," in *Proceeding of the 21st international scientific conference globalization and its Socio-Economic Consequences. SHS web of conferences*, 129, 09001 2021. doi: 10.1051/shsconf/202112909001
- Aulck, L., Velagapudi, N., Blumenstock, J., and West, J. (2016). Predicting student dropout in higher education. *arXiv [Preprint]* arXiv:1606.06364
- Balfanz, R., Herzog, L., and Mac Iver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educ. Psychol.* 42, 223–235. doi: 10.1080/00461520701621079
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Cameron, S. V., and Heckman, J. J. (2001). The dynamics of educational attainment for black, hispanic, and white males. *J. Polit. Econ.* 109, 455–499. doi: 10.1086/321014
- Chen, T., and Guestrin, C. (2016). "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, (San Francisco CA), 785–794. doi: 10.1145/2939672.2939785
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018
- Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inform. Theory* 13, 21–27. doi: 10.1109/TIT.1967.1053964
- Douzas, G., Bacao, F., and Last, F. (2018). Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Inform. Sci.* 465, 1–20. doi: 10.1016/j.ins.2018.06.056
- Duda, R. O., Hart, P. E., and Stork, D. G. (2001). *Pattern classification*, 2nd Edn. Hoboken, NJ: Wiley.

Acknowledgments

We would like to acknowledge the financial support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico, in the production of this work. We also would like to acknowledge Vicerrectoría de Investigación y Posgrado, the Advanced Artificial Intelligence Research Group, the Cyber Learning and Data Science Laboratory, and the Science Department of Tecnológico de Monterrey, Mexico City Campus, in the processes of this research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2023.1244686/full#supplementary-material>

- Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R., and Smink, J. (2008). *Dropout prevention: A practice guide (NCEE 2008-4025)*. Washington, DC: National Center for Education Evaluation and Regional Assistance.
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7, 179–188. doi: 10.1111/j.1469-1809.1936.tb02137.x
- Freund, Y., and Schapire, R. E. (1996). “Experiments with a new boosting algorithm,” in *Proceedings of the 13th international conference machine learning*, (San Francisco, CA), 148–156.
- Garg, A., Lilhore, U., Ghosh, P., et al. (2021). “Machine learning-based model for prediction of student’s performance in higher education,” in *Proceedings of the 8th international conference on signal processing and integrated networks, SPIN*, (Noida), 162–168.
- Gonzalez-Nucamendi, A., Noguez, J., Neri, L., Robledo-Rella, V., and García-Castelán, R. M. G. (2023). “Identifying the factors that affect higher education students at risk of dropping out IEEE,” in *Proceedings of the frontiers in education conference (FIE)*, (College Station, TX).
- Govender, C. M. (2020). Hopes, challenges and goals—voices of first-year at-risk higher education students in South Africa. *S. Afr. Rev. Sociol.* 51, 55–69. doi: 10.1080/21528586.2020.1806919
- Hedge, V., and Prageeth, P. P. (2018). “Higher education student dropout prediction and analysis through educational data mining,” in *Proceedings of the 2018 2nd international conference on inventive systems and control (ICISC)*, (Coimbatore), 694–699.
- Heublein, U. (2014). Student drop-out from german higher education institutions. *Eur. J. Educ.* 49, 497–513. doi: 10.1111/ejed.12097
- Hsu, C. W., and Yeh, C. C. (2019). Mining the student dropout in higher education. *J. Test. Eval.* 48, 4563–4575. doi: 10.1520/JTE20180021
- Lamb, S., Markussen, E., Teese, R., Sandberg, N., and Polesel, J. (eds.) (2010). *School dropout and completion: International comparative studies in theory and policy*. Berlin: Springer Science & Business Media, doi: 10.1007/978-90-481-9763-7
- Liu, T., Wang, C., Chang, L., and Gu, T. (2022). Predicting high-risk students using learning behavior. *Mathematics* 10:2483. doi: 10.3390/math10142483
- Nagy, M., and Molontay, R. (2018). “Predicting dropout in higher education based on secondary school performance,” in *Proceedings of the 2018 IEEE 22nd international conference on intelligent engineering systems (INES)*, (Las Palmas de Gran Canaria), doi: 10.1109/INES.2018.8523888
- Olaya, D., Vásquez, J., Maldonado, S., Miranda, J., and Verbeke, W. (2020). Uplift modeling for preventing student dropout in higher education. *Decis. Support Syst.* 134:113320. doi: 10.1016/j.dss.2020.113320
- Powers, D. M. (2020). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv [Preprint]* arXiv:2010.16061
- Quinlan, J. R. (1986). Induction of decision trees. *Mach. Learn.* 1, 81–106. doi: 10.1007/BF00116251
- Rodríguez Rojas, O. (2022). *Notas del curso, ML207 calibración de modelos en python del programa experto en machine learning de PROMIDAT generación Ada Lovelace*. Granadilla: PROMIDAT.
- Romesburg, C. (2004). *Cluster analysis for researchers*. Morrisville, NC: Lulu Press.
- Saravanan, T., Nagadeepa, N., and Mukunthan, B. (2022). “The effective learning approach to ICT-TPACK and prediction of the academic performance of students based on machine learning techniques,” in *Communication and intelligent systems. Lecture notes in networks and systems*, Vol. 461, eds H. Sharma, V. Shrivastava, K. Kumari Bharti, and L. Wang (Singapore: Springer), doi: 10.1007/978-981-19-2130-8_7
- Tec21 (2022). *Modelo Tec21*. Available online at: <https://tec.mx/en/model-tec21> (accessed September 8, 2022).
- Von Hippel, P. T., and Hofflinger, A. (2020). The data revolution comes to higher education: Identifying students at risk of dropout in Chile. *J. High. Educ. Policy Manag.* 43, 2–23. doi: 10.1080/1360080X.2020.1739800
- Wexler, J., and Pyle, N. (2012). Dropout prevention and the model-minority stereotype: Reflections from an Asian American high school dropout. *Urban Rev.* 44, 551–570. doi: 10.1007/s11256-012-0207-4



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Maria Rauschenberger,
University of Applied Sciences Emden Leer,
Germany
Marcia Håkansson Lindqvist,
Mid Sweden University, Sweden

*CORRESPONDENCE

Barthélémy Michalon
✉ bmichalon@tec.mx
Claudia Camacho-Zuñiga
✉ claudia.camacho@tec.mx

RECEIVED 01 July 2023

ACCEPTED 28 September 2023

PUBLISHED 13 October 2023

CITATION

Michalon B and Camacho-Zuñiga C (2023)
ChatGPT, a brand-new tool to strengthen
timeless competencies.
Front. Educ. 8:1251163.
doi: 10.3389/feduc.2023.1251163

COPYRIGHT

© 2023 Michalon and Camacho-Zuñiga. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

ChatGPT, a brand-new tool to strengthen timeless competencies

Barthélémy Michalon^{1*} and Claudia Camacho-Zuñiga^{2,3*}

¹School of Social Sciences and Government, Tecnológico de Monterrey, Puebla, Mexico, ²Institute for the Future of Education, Tecnológico de Monterrey, Monterrey, Mexico, ³School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico

In November 2022, the public release of ChatGPT, an artificial intelligence (AI)-based natural language model, was a groundbreaking point in many sectors of human life and education was not the exception. We describe how ChatGPT was integrated in an undergraduate course for an International Relations program in a private Mexican university. Under an action research methodology, we introduced this novel instrument in a course on Future Studies. Students were evaluated on their ability to explain to ChatGPT several discipline-specific methods and to make the AI implement these methods step by step. After six such activities, the outcomes evidenced that the students not only learned how to use the new AI tool and deepen their understanding of prospective methods, but also strengthened three soft or transversal competencies: communication, critical thinking, and logical and methodical reasoning. These results are promising in the framework of Skills for Industry 4.0 and Education for Sustainable Development; even more, they demonstrate how ChatGPT created an opportunity for the students to strengthen, and for the professor to assess, time-tested competencies. This is a call-to action for faculty and educational institutions to incorporate AI in their instructional design, not only to prepare our graduates for professional environments where they will collaborate with these technologies but also to enhance the quality and relevance of higher education in the digital age. Therefore, this work contributes to the growing body of research on how Artificial Intelligence (AI) can be used in higher education settings to enhance learning experiences and outcomes.

KEYWORDS

soft competencies, artificial intelligence in education, critical thinking, communication, educational innovation, higher education, organized reasoning

1. Introduction

With ChatGPT¹ now freely available to whoever wants to use it, certain classical tasks have suddenly become obsolete, if their completion can be fully or mostly delegated to this new Artificial Intelligence (AI) tool (Eloundou et al., 2023).

ChatGPT's integration into higher education has been received by the academic community with both enthusiasm and caution. On the positive side, ChatGPT offers a range of benefits that can significantly enhance the learning experience not only in higher education, but even toward self-directed learning and life-long learning. Notably, it provides real-time feedback and guidance

1 While GPT stands for "Generative Pre-trained Transformer," it has also been associated with "General Purpose Technology" (Eloundou et al., 2023).

to assist students in staying on track and addressing challenges as they arise. Additionally, its accessibility across various platforms, such as websites, smartphone apps, and messaging services, allows learners to engage with the tool at their convenience, fostering flexible learning. Furthermore, ChatGPT provides personalized support tailored to each learner's choices and goals and has the potential to enhance the use of open educational resources (Firat, 2023). This can be especially beneficial for self-directed learners and for learners who might not have access to traditional institutional education, therefore democratizing knowledge, making it more accessible to a broader audience.

Among the main concerns is the threat ChatGPT poses to traditional assessment methods. The tool's sophisticated text-generating capabilities can produce essays which raises questions about the integrity of student submissions and the validity of assessments, potentially undermining the very foundation of academic evaluation (Neumann et al., 2023; Rudolph et al., 2023). Additionally, ChatGPT can generate incorrect though seemingly relevant and accurate content, which can lead to a lack of critical thinking, if students overly rely on the tool for answers without deeply engaging with the content (Rudolph et al., 2023). Another issue is its fast-paced improvement, making it increasingly challenging for educational institutions to monitor or regulate its use.

In front of this new situation, it could be tempting to prevent the students from using ChatGPT or any other AI tool, nonetheless this path leads nowhere. Not only there is no way to enforce such a prohibition, but even if there were one, it would not make much sense to oblige the students to stay away from a type of tools whose use is destined to gradually become generalized—for good or for bad.

Hence, the need to choose a different direction, where we educators recognize the new existing situation. In the worst case we must adapt to it, while in the best scenario we can even take advantage of it. Concretely, in the short term, it means getting rid of certain tasks that have traditionally been assigned to students and embracing new ones.

This article represents an attempt to address this growing challenge: now that AI tools are able to complete a multiplicity of tasks almost instantaneously, how can these new instruments be used to foster and improve teaching and learning processes, instead of undermining them?

Our quest for an answer took the form of action research, based on an experiment carried out within the framework of a university course. First, various ways of incorporating ChatGPT as a pedagogical resource were explored and tested. Then the observation of the first set of results allowed us to select one of these ways. Finally, the merits of such a strategy were identified: it not only improved the students' technological literacy and understanding of key elements from the course itself, but also stimulated the acquisition of time-tested competencies. These competencies have proved to be useful in contexts that include and even transcend both the course and the ability to productively use this particular technological tool.

Consequently, this generates an apparent paradox: a vanguard instrument proves to be an effective way to practice and develop timeless competencies.

2. Materials and methods

This research is decidedly grounded on an empirical basis, provided by an in-class experiment in April–May, 2023, at the early stages of the

adoption of ChatGPT by university students. We evaluated several strategies for effectively integrating this emerging Artificial Intelligence (AI) tool as a teaching resource in the specific context of the lecture and eventually selected the one that, in our context, showed the most potential.

2.1. Onboarding a research initiative

In March, 2023, the Institute for the Future of Education (IFE) launched an invitation to professors to action research involving the insertion of AI tools in teaching environments.

Action research is a collaborative, reflective process where educators engage in systematic inquiry to improve their teaching practices and the learning experiences of their students. It involves a cyclical process of identifying a problem, planning a change, implementing the change, and then reflecting on the results to inform further action (Davison et al., 2004; Meyer et al., 2018; Voldby and Klein-Døssing, 2020). This iterative process, as one of its defining features, allows for continuous improvement and adaptation of the methodology based on the findings and reflections from each cycle. And, above all, this research approach emphasizes the active involvement of educators in studying and improving their own practices and sharing their results with the academic community (Davison et al., 2004; Meyer et al., 2018; Voldby and Klein-Døssing, 2020).

In response to this call from IFE, action research was deployed with the aim of identifying how ChatGPT's potential could be exploited at the university level. As many other educators also were at that time, the first author of this article was puzzled and intrigued by this AI tool recently released by OpenAI. The IFE initiative provided a framework to push boundaries and move forward into this new world. An ongoing course on Future Studies represented an adequate setting for such an exploration. Moreover, exploring what tomorrow might hold with the help of a tool that seemed to come from the future seemed to be a stimulating and elegant way to proceed.

The basic research protocol laid down by the IFE required to integrate the use of ChatGPT (Version 3.5) in at least four class activities or assignments, to register data throughout the whole process and finally to closely analyze the results from the experiment. This document constitutes the product of such an effort, in the pursuit of finding an answer to the research question exposed above.

2.2. The setting of the experiment

The experiment was to be conducted in the context of the English-taught course "Future Scenarios on the International Political Economy," taught at the Tecnológico de Monterrey, Campus Puebla in the Spring semester of 2023. A total of 19 students were initially enrolled in the course, all of whom were specializing in International Relations (IR). Approximately one-third of the students were pursuing a dual degree in IR and Economics, while another third of them were pursuing a dual degree in IR and Law. The students in the latter two groups were in their sixth semester of undergraduate studies, while the remaining students, who were solely enrolled in IR, were in their fourth semester. As a result, all students fell within the range of 19–21 years old.

We adhered to the following ethical considerations to ensure the integrity of our research and the protection of our participants' rights and privacy. All participants were over 18 years old, aligning with the definition of formal citizens under the Mexican legal framework. Prior to their involvement, each participant provided explicit consent to partake in the study. They were thoroughly informed about the research and confirmed their understanding of how their information would be used. It is paramount to note that all data gathered was treated with the utmost confidentiality, ensuring participant anonymity. Furthermore, the handling and storage of this data strictly adhered to the Privacy Policies set forth by the higher education institution overseeing the study.

On April 11th 2023, ChatGPT as a tool was introduced to the whole group. This initial session was meant to properly launch the experiment. First, some general and contextual information was provided about the software. In the meantime, the students without an OpenAI account² signed up for one. Following that, the professor offered a range of recommendations regarding the most effective approaches for engaging with ChatGPT, including: (1) Formulate the right prompts, (2) Interact repeatedly with it until you get what you want/need, (3) Be creative in what you ask it to do, including by using *personas*,³ (4) Take its answers critically, and (5) Use it as a complement of your own effort.

Next, two exercises were successively applied: groups of students were tasked first to instruct the AI to build a SWOT (Strengths, Weaknesses, Opportunities, and Threats) matrix on a given famous individual, and then to generate several multiple-choice questions about another method known as the Problem Tree. Importantly, both methods had been exposed in previous lessons, so the exercise was intended as a way to look again at the same contents through a different lens. Next, a class discussion took place on the outcomes generated in the process, with the aim of highlighting both mistakes and good practices.

2.3. Setting the baseline: students' prior knowledge and perceptions of ChatGPT

To close the launching session, all the students completed a survey regarding their own perceptions and experience (if any) about ChatGPT. The survey was designed in English, powered by Google Forms, and included seven multiple choice questions. One of the key questions sought to assess the level of prior knowledge each participant had about the software, with answers to be placed on a Likert scale from 0 ("Nothing at all") to 5 ("Already a lot"). As depicted in Figure 1, there was a notable disparity in that regard, but none of the surveyed individuals declared fully ignoring the tool, that had been released to the public half a year earlier. Among the respondents, 42% indicated possessing an intermediate level (Likert category: 3) of prior knowledge about it, 47% rated their familiarity as higher than that level, and only 11% positioned themselves below it.

Further questions revealed that slightly more than half of the group had never used ChatGPT for academic purposes before. Moreover, almost half the students considered that the activities conducted that day represented an effective way to verify and consolidate their knowledge about the two abovementioned methods that were the basis for both exercises. A positive but moderate correlation (with a Pearson coefficient of 0.52) was found between having prior knowledge of the tool and perceiving the activities implemented in class that day as effective. Only one quarter of the group selected one of the two options indicating that they would be using this tool often or very often in the future, with almost 60% expressing uncertainty about it. These results underscore a relatively high level of initial skepticism right after this first contact in the context of this specific course.

However, almost two-thirds of the surveyed individuals expressed confidence that ChatGPT would allow them to "learn more thanks to a more efficient use of [their] time," while most of the remaining students selected the option about "learn[ing] differently, but probably in the same amount as before."

If anything, the contrast between these positions—most are unsure whether they would use it but are nevertheless valuing positively its utility in the learning process—reveals the mixed feelings toward such an innovative tool, at the same time impressive and error-prone, that had suddenly popped-up in their lives as university students. The apparent inconsistency between the results can also be tentatively explained by the fear of being seen as embracing too openly the use of a tool whose compatibility with academic honesty and good practices was (and arguably still is) in question.

2.4. Initial objectives

At this stage, the objectives of the research were that the students: (1) acquired the ability to implement certain methods for exploring the future, using ChatGPT as an alternate and possibly effective teaching tool, and (2) developed new skills, which would enable them to take advantage and become familiar with ChatGPT. Undoubtedly, as the first publicly available AI tool operating under natural language instructions, this software was opening the way to a new setting where AI would gradually be found in most aspects of our everyday lives. In other words, it was about making sure that they would develop "good habits" from the early stages of their adoption of ChatGPT, so they would be comfortable when using other AI tools in the longer run. Even if it can be anticipated that the next generations of such instruments will be much more sophisticated, being an "early adopter" of a given system is commonly understood as a long-term advantage (Tobbin and Adjei, 2012), even if such a system evolves and get transformed over time.

This path leaves aside the assumption that the younger generations, for the sole merit of belonging to a particular age-range, are naturally able to smoothly and quickly adopt new technological tools and features. This purported quality, commonly expressed under the popular "digital natives" label, is severely questioned by research on the subject (Selwyn, 2009; Bennett and Maton, 2010; Margaryan et al., 2011). On the contrary, undergraduate students do not have an above-normal capacity to swiftly realize and learn how to properly use a new digital tool: Instead, they need to receive explanation and training, as any other person would.

² <https://chat.openai.com/>

³ Which means to instruct ChatGPT to play a particular role (expert, teacher, decision-maker, advisor, etc.) and/or to tailor its answer to a particular audience.

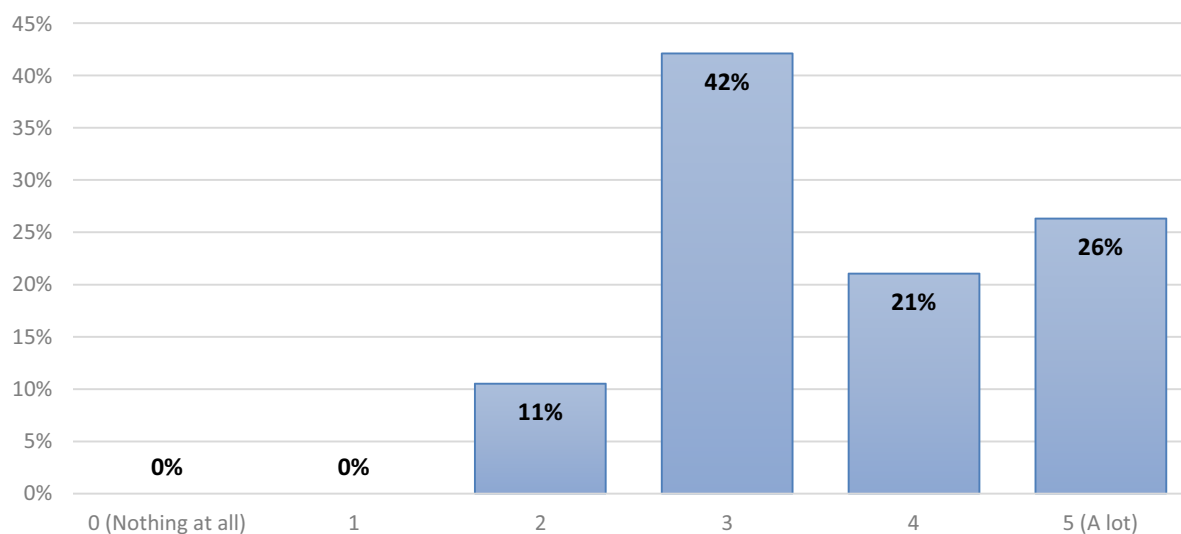


FIGURE 1

Prior knowledge of ChatGPT by the students ($n = 19$), according to their answers to: "What did you know about ChatGPT prior to today's [first] session?"

For these two motives, there was a compelling rationale for integrating exercises involving the use of ChatGPT into the course. Guided practice would facilitate a better assimilation of how the discipline-specific methods operate, and how ChatGPT functionalities can be productively and meaningfully mobilized.

2.5. The activities and their coding

Prior to its implementation, the experiment aimed at testing the relative merits of three possible uses of ChatGPT in the context of this course, whose content consisted in teaching a series of methods for developing a prospective project research.

It was initially assumed that this tool could be utilized by the students to perform one or several of the following functions, understood as the fulfillment of a particular task:

- (i) to summarize/reformulate known information about methods,
- (ii) to practically implement the methods that had already been studied in class and/or
- (iii) to discover and learn about methods that they did not know yet.

These three functions were selected for their potential to diversify and enrich the process through which students would learn and understand the functioning of prospective methods.

As detailed in Table 1, six different activities were applied in the context of this experiment, throughout four successive assignments. Each activity was assessed under three criteria: the quality of the prompts, the effectiveness / productivity of the interaction with ChatGPT, and finally, the extent to which the achieved outcome was fulfilling the requirements. Each student was given specific written feedback on their performance in each area but, for the purpose of this experimentation, each activity was eventually given an assessment in terms of "Positive," "Basic" or "Negative" for each of the three dimensions mentioned earlier. After being evaluated, each assignment

was discussed in the classroom, in order to highlight both the good practices and the most common missteps.

Notably, not all the activities were characterized by the same level of complexity.⁴ This discrepancy was not intentional but rather an inevitable outcome of conducting this experiment within real-life conditions, specifically within the context of a course that pursued its own pedagogical objectives. Therefore, the evolution of the students' success rate over time should be interpreted cautiously. However, even if the last task was the most challenging for being based on the most advanced prospective method, it was also the one where the performance of the students turned out to be the highest.

The deliverable consisted in a Word document, which included a link to a webpage where the whole interaction with ChatGPT would be shown. At the time of the experiment, the function to generate such a link was not embedded into the software, so we resorted to *Share GPT*, a Chrome extension specifically designed to that end.⁵

The rest of the same Word document served to display the student's answers to three or four questions to guide their personal reflections about their experience when using ChatGPT. The purpose was to monitor the evolution of their thoughts about their ability to use the AI software and their perception of its utility in their learning process.

2.6. Discarding the *summary/discovery* functions to focus on *implementation/practice*

The first assignment consisted of two activities, both centered on already known Future Studies methods, with the aim of helping the students prepare for an imminent exam. Activity 1.1 consisted in having ChatGPT recapitulate the steps to be followed in the

⁴ Table 1 integrates a column with an estimated difficulty score for each activity.

⁵ <https://sharegpt.com/>

TABLE 1 Description of activities implemented in a Future Studies course, using ChatGPT.

Act.	Delivery date	Summarized instructions given to the students	Difficulty level	Function or task
1.1	April 18, 2023	Ask ChatGPT to describe all the steps that must be followed when applying the SMIC or the Schwartz Axes methods.	2	(i) Summary
1.2	April 18, 2023	Ask ChatGPT to produce a MICMAC matrix on the issue of “consumerism” and to fill it.	2	(ii) Application
2.1	April 25, 2023	Ask ChatGPT to detail how the IGO method works. <i>Note: a basic description of it was provided to the students.</i>	1	(iii) Discovery
2.2	April 25, 2023	Generate an example of how the IGO method could be applied in a concrete situation.	1	(ii) Application
3	May 5, 2023	Reproduce the same activity we had in class on the Régnier Abacus method. Choose a different topic and have ChatGPT invent at least 10 experts and 10 specific issues/questions.	2	(ii) Application
4	May 28, 2023	Reproduce the same activity we had in class. Choose a different movie and complete the first three steps of the implementation of the MACTOR method.	3	(ii) Application

The difficulty score ranges from 1 (easiest) to 3 (hardest), relative to one another.

implementation process of either of two methods, thereby testing the function “to summarize/reformulate known information about methods” identified above as (i). Activity 1.2 required the students to instruct ChatGPT to apply another method, which related to function “to practically implement the methods that had already been studied in class” (ii). In both cases, the methods had been studied and implemented in previous sessions.

The results that were obtained through Activity 1.1 highlighted that ChatGPT was ill-suited to provide information about the content and functioning of any given method, because of its well-known intrinsic limitations in terms of factual veracity.

Furthermore, specialized webpages, that can easily be found via a classical web search, would provide complete and more reliable information on the matter, thereby reducing the utility of resorting to AI for that purpose. It is true that ChatGPT offers added value by providing explanations in terms that are better suited to the student and addressing follow-up questions on the same subject. However, this potential added value becomes moot when there exists so much uncertainty about the soundness of the information it delivers.

In contrast, Activity 1.2 delivered more encouraging results insofar as the observed limitations had to do with how the students used the tool, instead of being the consequence of intrinsic flaws from the tool itself.

The second assignment was also made of two activities, both of them based on the same prospective method. In this case, the students had not heard about this method beforehand, and Activity 2.1 consisted in having ChatGPT explain to them such method, which corresponds to function “to discover and learn about methods that they did not know yet,” defined earlier as (iii). The instructions contained a short description of it to help the students formulate their prompts. Activity 2.2 was designed to further test function (ii), so it was substantially similar to 1.2: once more, it consisted in implementing a method—in this case, the one that had been discovered thanks to Activity 2.1.

At first sight, Activity 2.1 appeared to have been completed in a satisfactory fashion, since all the process-related indicators went up (see Figures 2, 3). However, two problems were detected: first, it was *too simple* for the students to formulate proper prompts, since they only had to include in their instructions the information that they had received about the method. Consequently, this specific task did not require them to mobilize any reflection, creativity or effort of any kind. Second, the available information did not prevent some (minor) errors

from slipping into the generated answers, which exposed the students to inaccurate factual data about the method itself. In contrast, Activity 2.2 provided more interesting insights on how well the students were able to have ChatGPT follow their instructions and eventually produce the desired product.

The intermediate results obtained from these first four activities revealed that function (ii) “to implement the methods that had already been studied in class” was much more promising than the other two functions, (i) and (iii).

Consequently, the professor in charge of the experiment decided that, moving forward, the subsequent activities would no longer consist in the students using ChatGPT to summarize or discover methods, but exclusively in the students using it to *implement* such methods. Therefore, the next two tasks were designed to test and strengthen knowledge in a practical fashion, where each student would be giving instructions to ChatGPT about the steps to be followed in order to complete the full methodological process. In the context of this method-oriented course, it undoubtedly was a more promising way to take advantage of ChatGPT’s affordances.

3. Results

From an overall perspective, the experiment highlighted that despite a significant number of students rating their command of ChatGPT as decent or high, the learning curve proved to be steep: A total of six activities (in this case, distributed among four distinct installments) were indeed required to achieve that most students attain a satisfactory level of proficiency of ChatGPT.

This observation supports that the belief in the existence of “digital natives”—young people portrayed as being “innate, talented users of digital technologies” (Selwyn, 2009, p. 364)—is unfounded. Instead, a gradual adaptation and adoption process had to take place before most of the group proved to be able to make a fruitful use of the tool.

3.1. Prompts writing: a slow but continuous assimilation of the rules of engagement

A prompt is defined as “a phrase or question that is used to stimulate a response from ChatGPT” (Morales-Chan, 2023, p. 1). This

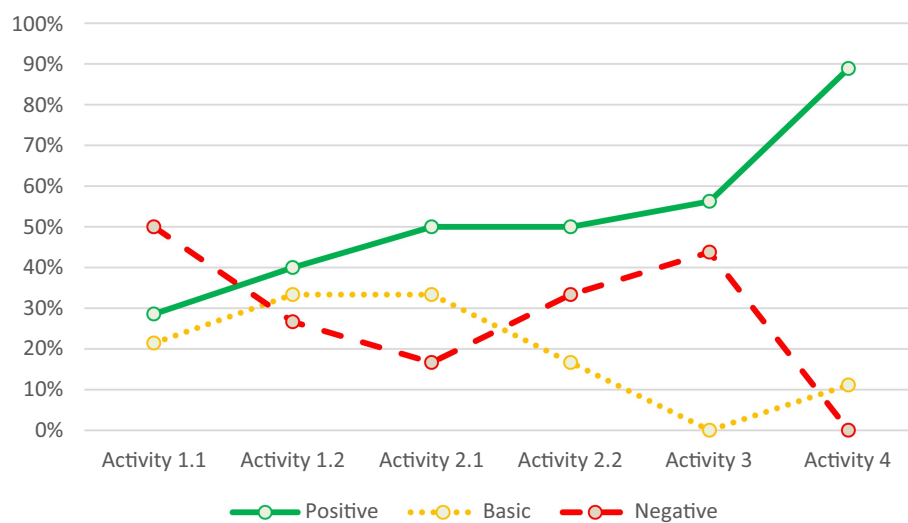


FIGURE 2

Percentage of students and the quality of their prompts per learning activity. The prompts were assessed positively, negatively, or in-between, by the professor (n = between 14 and 18, since not all of the students delivered every required assignment).

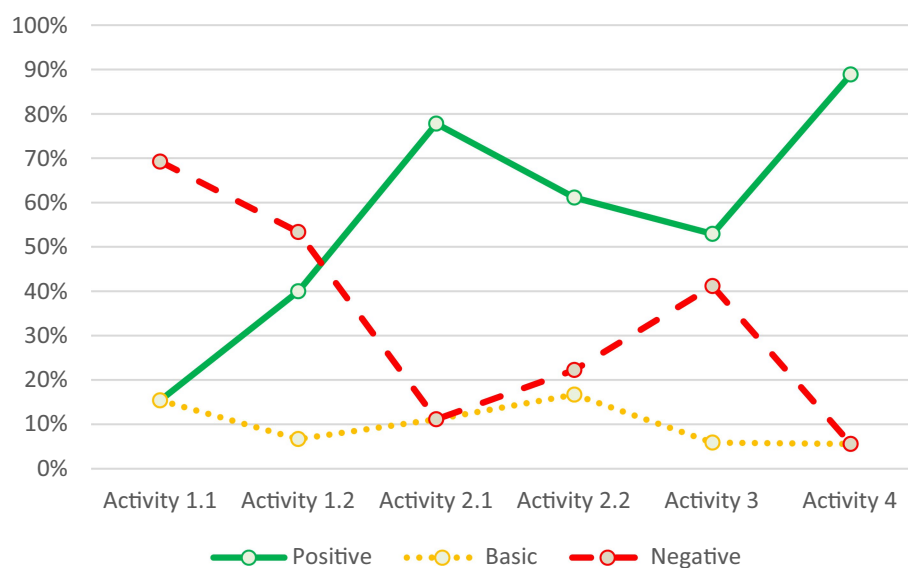


FIGURE 3

Percentage of students and quality of their “conversation” with ChatGPT per learning activity. The interaction was assessed by the professor as positively, negatively, or in-between (n = between 13 and 18, since not all of the students delivered every required assignment).

is, fundamentally, the human side of the interaction with AI. Consequently, the quality of the answer generated by the program heavily depends on the way the request was formulated in the first place (Morales-Chan, 2023, p. 2). This has become so clear that, at the beginning of 2023, training courses and articles on prompt writing—also known as “prompt engineering”—have become commonplace (see for instance an 18-h lesson offered on Coursera; White, 2023).

As shown in Figure 2, the quality of the prompts started from a low ground in the experiment, with less than a third of the students being able to properly explain to the machine what output they were expecting from it. At the beginning, even after taking an introductory session on the matter, a common mistake consisted in adopting the

same pattern as when using a search engine, with a string of keywords loosely connected with one another.

At later stages of the experiment, full sentences were finally becoming the regular practice: in Activities 2.1, 2.2, and 3, a narrow majority of students wrote adequate prompts, but the rest of them still lacked clarity and/or specificity. Activity 4, which was the final one, eventually showed an uptick in progress in that respect, with 90% of the group properly communicating with the AI tool. The rest of the group achieved a level that was considered basic, and no one showed plain incapacity to perform this task (see Figure 2).

Therefore, the trial-and-error process, conducted throughout six successive activities interspersed with feedback, allowed for a

significant improvement in the proportion of students capable of “communicating” through clear and specific prompts.

3.2. The quality of the “conversation” with ChatGPT: a rocky but upward trend

Unlike using search engines, engaging with ChatGPT involves a dynamic back-and-forth exchange between the user and the machine. Instead of a simple request and the presentation of a list of results, what unfolds is akin to a genuine interaction or conversation: This is precisely the reason why the word “chat” that was appended to GPT, the latter referring to the underlying technology that powers it (Eloundou et al., 2023).

Familiar with digital tools that typically involve one-off queries (even if some of them might end up being repeated under different forms as part of the same research process), many students initially lacked the inclination to engage in a continuous exchange with ChatGPT. They did not readily recognize the opportunity, or even in certain cases the need, to follow a step-by-step interaction, to request more details, or even highlight mistakes and ask for rectifications.

Engaging with ChatGPT entails more than just writing additional prompts after the initial one: it also implies analyzing and understanding the generated answers and tailoring subsequent prompts accordingly. In the context of the activities based on function (ii) “to practically implement the methods that had already been studied in class,” it further assumes that the user possesses sufficient clarity about how the entire process can be broken down, and what outcomes are expected at each stage.

In line with the earlier observations regarding prompts writing, it also took some time before many of the students were able to conduct a meaningful and productive interaction with the tool. In particular, they were prone to simply accept what the program delivered to them and moved from there without further questioning the generated statements.

As shown in Figure 3, the quality of the interaction with the software showed an uneven but noticeable progress, starting below 20% and ending close to 90%. The ups and downs in-between can be explained by the varying levels of difficulty of the different methods: a more demanding method would require a longer and arguably more sophisticated discussion with ChatGPT, which raises the likelihood that students commit errors in the process. However, the fact that the last—and most challenging—method was the one where, on average, the best interactions were observed is compelling evidence that this point ended up being properly assimilated by most of the group.

3.3. Generated outcomes: a good process is a necessary but not a sufficient condition for positive outputs

After focusing on two dimensions based on the *process* of using ChatGPT (sections 3.1 and 3.2), the third criterion had to do with how effective this interaction had been, as revealed by the quality of the *product* that was eventually generated by ChatGPT when following each student’s instructions.

Figure 4 reveals that almost all the outcomes generated for Activity 1.1 were substandard. Regardless of which specific method they chose (they had to pick one, among two that had previously been

studied in class), almost all the students gathered incomplete and/or false data about it. They proved unable to deal with the discrepancy between what they had learnt earlier in class on this subject, and what ChatGPT was—misleadingly—delivering to them.

While the software generated answers with inaccurate and fragmented information about the method, it presented them with such apparent confidence that the students simply accepted this alternate—and fictitious—version of what the method consisted of.

The outcomes for Activity 1.2 (which, as a reminder, was delivered at the same time as Activity 1.1) were significantly better since close to 25% of the group got it right. Since the methods to be used in both parts of the first assignment (i.e., Activities 1.1 and 1.2) were equally challenging in terms of intrinsic complexity, the explanation for the gap between the observed performances in each case had to be found elsewhere, namely in the kind of exercise faced by the students. In this case, ChatGPT’s propensity to invent to fill in knowledge gaps was not at the fore, since the AI program was merely asked to implement a series of steps. Therefore, the quality of the eventual outcomes did not depend that much on ChatGPT’s being truthful and/or accurate, but primarily on each student’s capacity to properly use and “guide” the tool throughout the different phases that constitute each method. This observation stood as a confirmation that function (ii) was more promising than the other two, since the latter rely too heavily on ChatGPT’s changing capacity to stick to real information.⁶

In the class discussion that followed, further emphasis was made on the software’s limitations in terms of factual veracity, so the students were encouraged not to settle too easily with ChatGPT’s first answers.

The next three activities delivered better outcomes, with between 50 and 60% of the students managing to obtain satisfactory answers from the software (see Figure 4). Despite being a majority, it can still be seen as relatively low percentages, all the more so since they did not show significant increase over time. The last activity even showed a decline in the rate of positive outcomes, which can be explained by the more challenging nature of the exercise, based on the most complex method. Even if only one third of the students had it right, almost everyone else ended up obtaining a “basic” outcome, with only one of them reaching a final result assessed as “negative.” This latest observation allows to relate the drop in the “positive” curve with the increased difficulty of the task and therefore to relativize the significance of such a break in the upward trend.

It is worth noticing that a higher rate of “positive” assessments in terms of process usually translates into more “positive” assessments in terms of outcomes, and the same applies for negative assessments and outcomes. As illustrated by the proximity between the curves on

⁶ It has to be conceded, however, that function (i) (generating a summary of about an already-known information) offers a way to use at our advantage ChatGPT’s tendency to present its own creations as facts, as long as the students are instructed to detect mistakes in the generated answers and to have them rectified. Nonetheless, this option remains viable only when ChatGPT’s mistakes are confined to specific aspects of an otherwise well-understood method. In our experience, the program often resorted to fully inventing a given method that it did not know beforehand, starting from the scant elements that were provided to it. This adaptation strategy from the program renders the task to correct it not only overwhelming but also of limited utility for learning purposes.

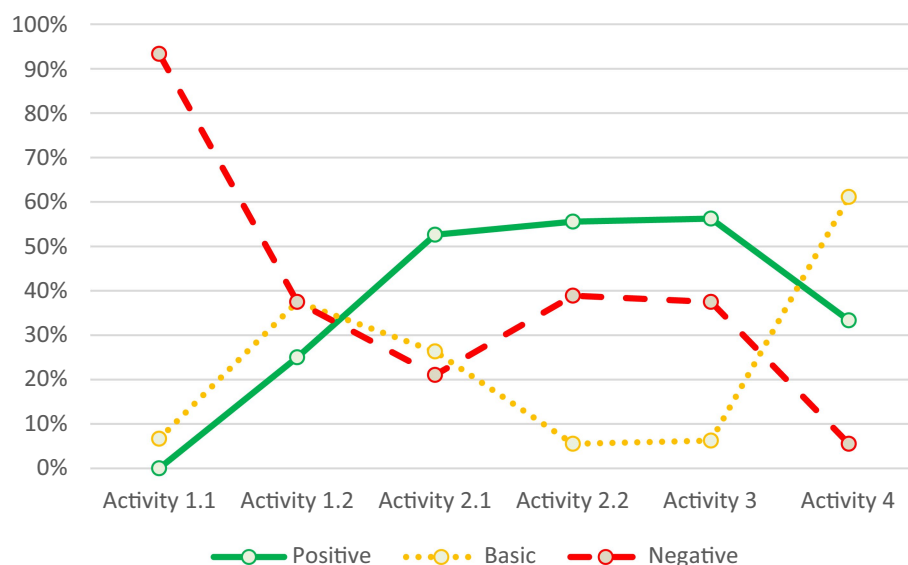


FIGURE 4

Percentage of students and quality of the outcome generated by ChatGPT per learning activity. The outcomes assessed by the professor as positively, negatively, or in-between (n = between 15 and 19, since not all of the students delivered every required assignment).

Figures 5, 6, the rates are closely correlated, except for the first two activities and the very last one. In the first two cases, which were both part of the first assignment, this can be explained by the fact that this was the first time the students were completing a ChatGPT-based homework and were only starting to navigate the learning curve.

In the last case, the significant increase of the method's complexity is an adequate explanatory factor: even if almost all the students (close to 90%) had performed very well in terms of process, it was no guarantee that the outcome would be satisfactory in the same proportions. Indeed, even if the prompts are formulated within the recommended parameters and if the conversation is effectively sustained, the final result might be judged as falling short of the expectations, if some confusions or misunderstandings subsisted in the students' minds regarding how exactly the method is supposed to be operated.

However, this by no means indicates that the activity itself was a failure: on the contrary, it played a role as a powerful and straightforward indicator of which specific parts of the method had not been properly understood by the students.⁷ This new information offered the opportunity for a precisely targeted complementary explanation and/or exercise during the next session.

3.4. How do students self-assess their own progress at using ChatGPT?

In addition to the activities themselves, each of the four assignments also consisted in each student providing a "final reflection" about their experience when dealing with ChatGPT.

As a confirmation of the "learning curve" that still had to be navigated in the first assignment, 87% of the students answered that

it had been "hard" to complete it by using ChatGPT, while only 6% qualified the task as "easy," and the same proportion provided a mixed answer⁸ (see Figure 7). The answers to the same question about the second assignment (on Activities 2.1 and 2.2) revealed a radical switch in perceptions, since this time only 11% selected "hard" and 61% "easy," with the rest sitting on the fence.⁹

For the third assignment (Activity 3), the question was modified to focus on their self-assessed ability to use the tool,¹⁰ and close to 100% of the respondents considered that they did actually improve it, expressing confidence that they were now proficient in handling the tool. However, these numbers were much higher than the percentages of students who had managed to properly apply the processes. This gap reveals that many students—almost half of them—were taking for granted that they were good enough using ChatGPT, while it was in fact not the case yet. In contrast, self-assessments and external evaluation became much more aligned with one another for the last assignment (Activity 4).

This transitory gap between actual performance and its perception was also noticeable through their answers to a question, asked as part of each of the four assignments, about the "techniques" that they had applied in their interaction with ChatGPT. Since the very first assignment, the students had been consistently providing relevant and correct advice, by describing good practices such as being specific in the requests, providing relevant information in the prompt itself, being patient and organized in the interaction, checking the accuracy of the generated contents, or signaling possible mistakes. However,

8 The exact question was "How hard was it for you to get what you wanted from ChatGPT?"

9 For the purpose of building Figure 7, the answer "hard to complete it by using ChatGPT" was eventually recoded as a "low" capacity to use ChatGPT and, conversely, "easy" was recoded as a "high" ability.

10 "Do you feel that you have improved your capacity to get what you wanted from ChatGPT?"

7 As evidenced by their incomplete or flawed instructions and/or their failure to identify inaccurate or contradictory elements in the generated answers.

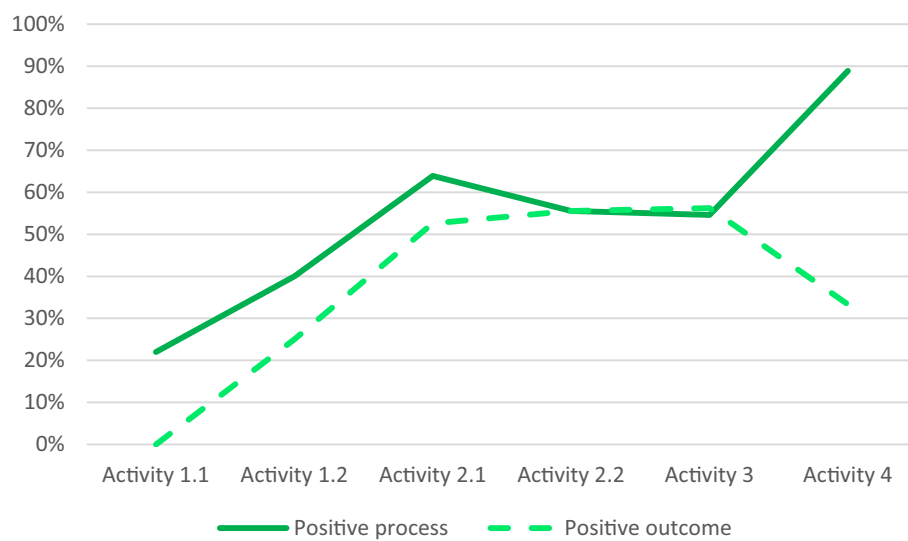


FIGURE 5

Percentages of positive assessment of outcomes and positive assessments of the interactions with ChatGPT (calculated as the average rate of positive assessments of prompts and conversations with the tool) per learning activity (n = between 15 and 19, since not all of the students delivered every required assignment).



FIGURE 6

Percentages of negative assessment of outcomes and negative assessments of the interactions with ChatGPT (calculated as the average rate of negative assessments of prompts and conversations with the tool) per learning activity (n = between 15 and 19, since not all of the students delivered all the required assignments).

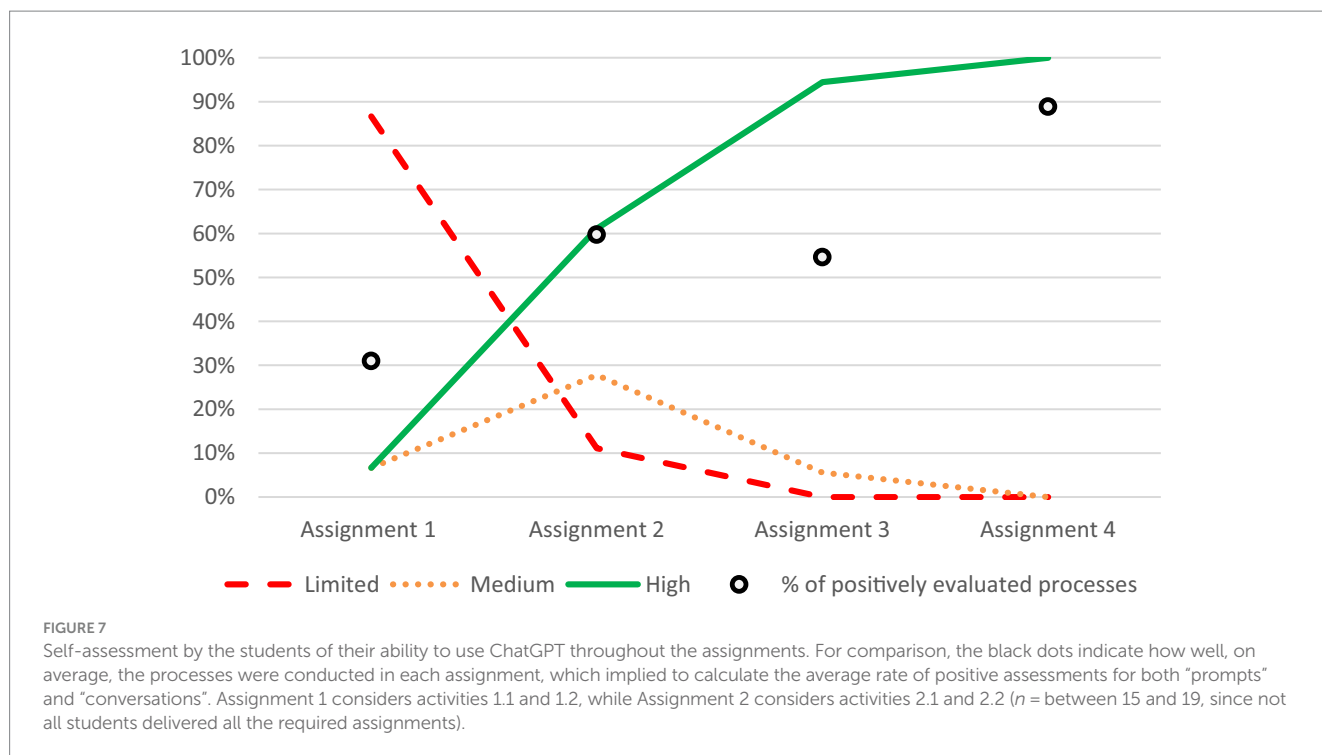
except for the very last assignment, there was a notable delay between the correct verbalization of the need to proceed in a particular way, and the actual implementation of it.

This latest observation further reinforces the point that it is not enough to inform the students about how to proceed, or to hear or read them describe the process (Reese, 2011). It is rather much more fruitful to have them *actually do it*, which allows the professor to observe to what extent their own practice is consistent with what they state should be done. In fact, it has been shown that learning-by-doing plays a key role for digital literacy since students learn how to use technology by using it (Tan and Kim, 2015).

4. Discussion

Even though the experiment was conducted with two objectives in mind (teaching prospective methods through novel means and providing the students with an introduction to the use of ChatGPT), it turned out that this initiative produced another, unexpected, output.

Indeed, over the course of the experiment, the professor noted that, in order to properly execute the ChatGPT-based tasks, the students had to mobilize certain competencies that can be qualified as “classical” or “timeless,” since they have long been identified as key within any student’s formation. When sharing feedback, it soon



became clear that the central conditions for reaching their objectives when interacting with ChatGPT did not have to do with ChatGPT-specific requirements, but with being able to mobilize three competencies with a much broader range and relevance: written expression, critical thinking, and methodical and logical reasoning. These competencies are identified as soft, transversal, or non-cognitive, to reflect that they are the basis for a successful performance within the professional and personal spheres (OECD, 2015).

Consequently, using ChatGPT appeared to bear another merit: not only it would contribute to the course's own objectives in terms of content, and help students assimilate how to handle a new digital tool, but it would at the same time stimulate the development of competencies that are widely recognized as being essential for any (future) professional, especially (but of course not exclusively) in the field of social sciences.

Additionally, ChatGPT proved to serve as an efficient revelator of the extent to which each student has achieved to assimilate each of these competencies, which is particularly useful from the educator's perspective.

4.1. Written communication: practicing and improving through prompt formulation

Given its capacity to generate text with correct grammar, spelling, and logic at a pace faster than it can be read out loud, ChatGPT may understandably be perceived as a potential threat to the development of writing skills if individuals start relying heavily on it for this task. However, ChatGPT also presents opportunities in this regard, since it necessitates the formulation of precise and clear prompts to obtain outputs that align with initial expectations. As a result, it can serve as a catalyst for

enhancing the ability to articulate thoughts effectively and communicate intentions with clarity.

For the Tecnológico de Monterrey, communication as a competency refers to effectively use different languages, resources, and communication strategies according to the context, in their interaction within various professional and personal networks, with different purposes or objectives (Olivares et al., 2021). It is a transversal competency that is relevant for successful performance within the professional and personal environment (OECD, 2015). It is understood as comprising language proficiency, presentation competency, capacity for dialogue, communication readiness, consensus orientation and openness toward criticism (Ehlers, 2020). Therefore, the focus is on information purposes as well as strategic communication skills in order to be able to successfully and appropriately deal with different contexts and situations. As a modality, written communication enables the student to express through writing ideas, arguments, and emotions with linguistic correctness and considering contextual elements, both in the mother tongue and in an additional language (Olivares et al., 2021).

Before and throughout the experiment, several tips were provided to the students as to what effective prompts look like. At some point, it seemed that all of them could be subsumed into one, that consisted in *writing as if you were communicating with a fellow human*. While it appears to be straightforward on its face, this advice nevertheless proved to be of little effectiveness, since today's students also tend to simply be as evasive and vague in their written exchanges with their “fellow humans.” It was therefore necessary to specifically insist, in the feedback sessions, on what constituted a good prompt, as opposed to an ineffective one.

In doing so, it was particularly helpful to point at the patent connection between an inadequately worded prompt and the unsatisfactory nature of the output generated by ChatGPT, thereby instantly revealing that “something is wrong.” This feedback presents

similarities with the one that a developer receives after committing mistakes in their programming process: since their code is not working as intended, something must be corrected somewhere. Before the public release of ChatGPT, this was not something that could be highlighted so clearly for natural language. The closest equivalent would be a tutor showing to the tutee, on a one-on-one interaction, that the sentences he or she has written are confusing or missing important points. By definition, such a process would be extremely hard to scale and would in any case be tributary to other parameters, such as the level of mutual understanding between the people involved. In contrast, ChatGPT offers promising prospects in this regard.

This is not to suggest that the students' writing experience should henceforth be limited to the formulation of instructions to an AI machine. The point here is that the interaction with ChatGPT offers opportunities to develop and strengthen this skill, in such a way that it can afterwards be exploited in other contexts.

4.2. Critical thinking: spotting ChatGPT's approximations and hallucinations

False or fake information is the well-known drawback of the trove of data available on the Internet. This risk is compounded on (and by) ChatGPT, first because it does not provide its own sources (whose reliability or lack thereof usually stands for a useful indicator for the cautious Internet user), and second because of its own self-confidence when generating outputs, which makes its statements appear sounder than they actually are.

Paradoxically, one of ChatGPT's main flaws—its very flexible relationship with truth—turns out to be a useful quality in the educational context since it represents a strong incentive for the students to double-check the generated outputs. Therefore, the mere fact of using it represents a powerfully illustrative case-study of the importance of not trusting blindly the information that has been generated. On the contrary, with every use of ChatGPT, students are reminded that they must refrain from taking the veracity of ChatGPT's answers for granted and must instead approach them with systematic skepticism, by applying their critical thinking. As denoted by the easy and quick dissemination of disinformation,¹¹ this reflex is not well established yet.

For the Tecnológico de Monterrey, a student with critical thinking evaluates the solidity of one's own and others' reasoning, based on the identification of fallacies and contradictions that allow forming a personal judgment in the face of a situation or problem (Olivares et al., 2021). Critical thinking allows questioning and changing perspectives in relation to existing identified facts (Ehlers, 2020); it is therefore related to other reasoning skills, like self-reflection and problem-solving competencies. Besides, it is a relevant competency in the education for sustainable development framework and it is essential for facing the threat of employment disruptions from automation and AI (PWC, 2020).

In the experiment, students played a leading role, which was intended to give them more confidence when it came to questioning the generated outcome: instead of simply querying for information

and passively receiving (and accepting) it, they were placed in a situation where they were instructed to gradually implement a given method and getting outputs in return. Since they had previously learnt this method in class, they were, at least theoretically, sufficiently equipped to critically examine and assess the value of the generated answers.

This context was designed as a means to test and practice critical thinking, defined as "skillful, responsible thinking that facilitates good judgment because it (1) relies upon criteria, (2) is self-correcting, and (3) is sensitive to context" (Lipman, 1988, p. 145). In the exercise, each student had to contrast what they had obtained with what they knew, first to identify if there existed a discrepancy between the two and, if any, find out whether it was the result of a flawed knowledge (or a wrongly worded prompt) on their side, or a misstep on ChatGPT's.

As shown in the results section, the critical thinking competency is the one that students are taking longer to assimilate, since the outcomes that most students had reached were still containing undetected errors, which is the sign of a premature and undeserved acceptance of the responses provided by ChatGPT.

4.3. Methodical and logical reasoning: organizing a thought process by sustaining a "conversation" with ChatGPT

To be meaningful, the two competencies highlighted in sections 4.1 and 4.2 have to be combined with a third one, which is the ability to structure the process in such a way that the different steps eventually lead to the desired outcome. First, the possible detection of factual errors or omissions in the generated responses serves as an opportunity to engage in a constructive dialogue with the tool, which aims to highlight and address these inaccuracies, ultimately leading to their rectification.

Second, carefully worded prompts would be of little use if they were not inserted in a logically articulated framework. In fact, it can even be argued that being able to correctly organize ideas (or, for this purpose, the successive steps) together is part of the written expression competency. In this article, they have been presented as separate to take due account of the fact that the described activities, designed around function (ii), fit neatly within a key feature of ChatGPT, which is its conversational dimension.

For the Tecnológico de Monterrey, the ability to solve problems and questions using logical and methodical reasoning in the analysis of clearly structured situations represents an incipient or basic level of "scientific thinking." This competency involves using structured methods in the analysis of complex situations from disciplinary and multidisciplinary perspectives and incorporating evidence-based professional practice (Olivares et al., 2021). Scientific thinking involves higher-order reasoning skills like analysis, evaluation, and synthesis of information (Suciati et al., 2018). As mentioned before, this competency relates with the ability to understand and solve complex problems (Vázquez-Parra et al., 2022).

Indeed, implementing different prospective methods, as mandated in four of the six activities, demands a comprehensive understanding of how each of them breaks down into a series of successive steps. Since ChatGPT allows for a sustained interaction between the user and the program, in a way that mimics human interaction, it allows the students to organize a communication process on this matter.

¹¹ Contrary to common perceptions, young people are not better protected against disinformation than older individuals (Pan et al., 2021).

Importantly, it also grants the professor an opportunity to observe and assess how it has been sequenced and conducted.

This organized reasoning can be assimilated to an improved “self-explanation,” defined as “prompting students to explain concepts to themselves during initial learning” (MIT Teaching + Learning Lab, 2023), insofar as they are indeed describing the successive steps of a given method, with enough details for these to be implemented. Technically, they are not explaining it to anyone else, but they do obtain a response from the AI tool, in such a way that they can assess the accuracy and completeness of their own input, and eventually make (or request) adjustments in the following stages of the “conversation.” It triggers a process that has its own merits in comparison to pure self-explanation (where no external feedback is delivered) and to an explanation actually directed to another person. In addition to being contingent on the other person’s own availability and ability to deal with the issue at stake, this second scenario is unlikely to place the student in a leading role, but rather in a position where they intend to guide someone who actually knows more about the process and will judge or correct them in case of missteps.

4.4. Toward an application of activities of this kind beyond methodology courses

It might appear that the description of this experiment only leaves a narrow margin of application for this specific use of ChatGPT, limited to courses dedicated to explaining how to use certain methods. This perception is not accurate, since this strategy can be transposed to any setting where students are expected to follow and implement a series of steps. The described exercise, which by no means forbids the implementation of additional, complementary activities where AI is not mobilized,¹² fundamentally consists in giving them the task to *lead* a process (with ChatGPT in the executing role) instead of merely *following* its steps.

Such logic can be applied in different contexts. For instance, students might be tasked with having ChatGPT writing a story or an essay, with the specific instructions that the AI tool has to be guided step by step throughout the process. As in the case of the described experiment, the evaluation should not concentrate—at least not primarily—on the generated outcome, but rather on the way the student guided the whole process, including by requiring corrections and improvements along the way. Hence the importance of including evidence of the interaction sequence, as exposed in section 2.5.

On a more general note, a key practical implication of this paper is that ChatGPT provides an opportunity for the professor to closely and directly observe how the student manages a complex process, instead of merely inferring from the outcome how well the process had been conducted. In order not to miss such an opportunity, the design and instructions of assignments should be adjusted with this extended target in mind.

¹² Which indeed applied in the context of the prospective course, where students had to apply the methods themselves both before the experiment (while they were taught about them) and after it (by integrating them in their respective final prospective research projects).

4.5. Limitations and future work

We recognize that, as an exploratory action research on a novel and quickly evolving groundbreaking tool, our research has limitations, which could be addressed in future work.

Our results are based on the professor’s evaluation of students’ interactions with, and outcomes from, ChatGPT as well as the students’ perception of their own ability to use this AI. For a more robust support for the findings in future study, researchers should consider the use of an evaluation instrument that would not be based only on the specific observable skills when handling ChatGPT (prompts writing, quality of the interaction and generated outcome) but directly on the competencies that this exercise inadvertently allowed to both test and foster (written expression, critical thinking, and logical and methodical reasoning).

Besides, the participants in this research were unfamiliar with the use of AI for learning applications. In the future, researchers will be dealing with students with more experience in the use of ChatGPT and other technological tools. This context is likely to require an adjustment to the instructions and expectations. For instance, most participants might be more comfortable in prompt writing, in comparison to the sample observed in the context of this experiment.

While our activities relied on using open access ChatGPT Version 3.5, other versions and even other apps and software have emerged since, offering more opportunities for supporting the learning process. So, future studies should analyze their own potential and/or actual integration into higher education.

It should also be noted that our research took place in the context of a private institution, where the students have access to computers, smartphones, and internet connection in a higher proportion than most of Latin American students. In this continent, digital divide has been evidenced as a limitation of population to access the internet and other information and communication technologies. Future research should consider collaborating with different universities for implementation of similar activities, to have a broader sample of Latin American students.

Finally, this longitudinal research was developed over a relatively short period: its temporal scope could be extended, either by replicating a similar experiment to a new generation of students, or by applying a more advanced experiment on the same initial group of participants.

5. Conclusion

Initially geared toward the transmission—and hopefully the eventual acquisition—of techniques to handle a novel digital tool, this experiment ended up providing additional arguments for the continuing relevance of competencies whose importance had been emphasized long ago: written expression, critical thinking, and organized reasoning.

This unexpected finding serves as confirmation that, despite the continuous evolution of our tools toward increased sophistication, there are certain skills and competencies that remain as essential as ever. Rather than diminishing their significance, AI technology is, in fact, strengthening the case for persistently fostering their development.

Data availability statement

The datasets presented in this article are not readily available because the participants agreed to participate under the condition of confidentiality and anonymity of their answers. Requests to access the datasets should be directed to bmichalon@tec.mx.

Ethics statement

All participants were over 18 years old, aligning with the definition of formal citizens under the Mexican legal framework. Prior to their involvement, each participant provided explicit consent to partake in the study. They were thoroughly informed about the research and confirmed their understanding of how their information would be used. It is paramount to note that all data gathered was treated with the utmost confidentiality, ensuring participant anonymity. Furthermore, the handling and storage of this data strictly adhered to the Privacy Policies set forth by the higher education institution overseeing the study. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

BM and CC-Z contributed to the conception and design of the study, identified and adapted the references and the theoretical

framework, and wrote the final manuscript. BM conducted the instructional design and its implementation, conducted the data collection and analysis, as well as, visualization of the results, and wrote the first manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors would like to acknowledge the support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico in the production of this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Bennett, S., and Maton, K. (2010). Beyond the 'digital natives' debate: towards a more nuanced understanding of students' technology experiences: beyond the 'digital natives' debate. *J. Comput. Assist. Learn.* 26, 321–331. doi: 10.1111/j.1365-2729.2010.00360.x
- Davison, R., Martinson, M. G., and Kock, N. (2004). Principles of canonical action research. *Inf. Syst. J.* 14, 65–86. doi: 10.1111/j.1365-2575.2004.00162.x
- Ehlers, U. D. (2020). *Future skills: the future of learning and higher education* Books on Demand Available at: <https://books.google.com.mx/books?id=lnTgDwAAQBAJ>.
- Eloundou, T., Manning, S., Mishkin, P., and Rock, Y. D. (2023). GPTs are GPTs: an early look at the labor market impact potential of large language models. *arXiv*. doi: 10.48550/ARXIV.2303.10130
- Firat, M. (2023). *How Chat GPT Can Transform Autodidactic Experiences and Open Education?* OSF Preprints. doi: 10.31219/osf.io/9ge8m
- Lipman, M. (1988). Critical thinking: what can it be? *Contemporary issues in curriculum*, Ornstein A. de and L. Behar. Boston, MA: Allyn & Bacon. Available at: <https://eric.ed.gov/?id=ED352326>.
- Margaryan, A., Littlejohn, A., and Vojt, G. (2011). Are digital natives a myth or reality? University students' use of digital technologies. *Comput. Educ.* 56, 429–440. doi: 10.1016/j.compedu.2010.09.004
- Meyer, M. A., Hendricks, M., Newman, G. D., Masterson, J. H., Cooper, J. T., Sansom, G., et al. (2018). Participatory action research: tools for disaster resilience education. *Int. J. Disast. Resil. Built Environ.* 9, 402–419. doi: 10.1108/IJDRBE-02-2017-0015
- MIT Teaching + Learning Lab. (2023). *Help students retain, organize and integrate knowledge*. Available at: <https://tll.mit.edu/teaching-resources/how-to-teach/help-students-retain-organize-and-integrate-knowledge/>.
- Morales-Chan, Miguel Ángel. (2023). *Explorando el potencial de Chat GPT: una clasificación de Prompts efectivos para la enseñanza*. Paper GES. Available at: <http://biblioteca.galileo.edu/tesario/handle/123456789/1348>.
- Neumann, M., Rauschenberger, M., and Schön, E.-M. (2023). "We need to talk about ChatGPT: the future of AI and higher education" in *In 2023 IEEE/ACM 5th International Workshop on Software Engineering Education for the Next Generation (SEENG)* (Melbourne: IEEE), 29–32.
- OECD (2015). *Skills for Social Progress: The Power of Social and Emotional Skills*. OECD Skills Studies, OECD Publishing, Paris. doi: 10.1787/9789264226159-en
- Olivares, S. L. O., Islas, J. R. L., Garin, M. J. P., Chapa, J. A. R., Hernández, C. H. A., and Ortega, L. O. P. (2021). *Modelo Educativo Tec21: retos para una vivencia que transforma* Tecnológico de Monterrey Available at: <https://books.google.com.mx/books?id=aF47EAAAQBAJ>.
- Pan, W., Liu, D., and Fang, J. (2021). An examination of factors contributing to the acceptance of online health misinformation. *Front. Psychol.* 12:630268. doi: 10.3389/fpsyg.2021.630268
- PWC. (2020). *Skills for industry curriculum guidelines 4.0: future proof education and training for manufacturing in Europe: final report*. Lu: Publications Office.
- Reese, H. W. (2011). The learning-by-doing principle. *Behav. Dev. Bull.* 17, 1–19. doi: 10.1037/h0100597
- Rudolph, J., Tan, S., and Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *J. Appl. Learn. Teach.* 6, 342–363. doi: 10.37074/jalt.2023.6.1.9
- Selwyn, N. (2009). The digital native – myth and reality. *ASLIB Proc.* 61, 364–379. doi: 10.1108/00012530910973776
- Suciati, S., Ali, M., Anggraini, A., and Dermawan, Z. (2018). The profile of XI grade students' scientific thinking abilities on scientific approach implementation. *J. Pendi. IPA Indon.* 7, 341–346. doi: 10.15294/jpii.v7i3.15382
- Tan, L., and Kim, B. (2015). "Learning by doing in the digital media age" in *New media and learning in the 21st century*. eds. T.-B. Lin, V. Chen and C. S. Chai (Singapore: Springer Singapore), 181–197. Education Innovation Series.
- Tobbin, P., and Adjei, J. (2012). Understanding the characteristics of early and late adopters of technology: the case of mobile money. *Int. J. E-Serv. Mob. Appl.* 4, 37–54. doi: 10.4018/jesma.2012040103
- Vázquez-Parra, J. C., Castillo-Martínez, I. M., Ramírez-Montoya, M. S., and Millán, A. (2022). Development of the perception of achievement of complex thinking: a disciplinary approach in a Latin American student population. *Educ. Sci.* 12:289. doi: 10.3390/educsci12050289
- Voldby, C. R., and Klein-Døssing, R. (2020). I thought we were supposed to learn how to become better coaches: developing coach education through action research. *Educ. Act. Res.* 28, 534–553. doi: 10.1080/09650792.2019.1605920
- White, J. (2023). *Prompt engineering for ChatGPT* Coursera Available at: <https://www.coursera.org/learn/prompt-engineering>.



OPEN ACCESS

EDITED BY

Kingsley Okoye,
Monterrey Institute of Technology and Higher
Education (ITESM), Mexico

REVIEWED BY

James Harold Davenport,
University of Bath, United Kingdom
Yousef Wardat,
Higher Colleges of Technology,
United Arab Emirates

*CORRESPONDENCE

Dadan Dasari
✉ dadan.dasari@upi.edu

RECEIVED 16 September 2023

ACCEPTED 29 December 2023

PUBLISHED 11 January 2024

CITATION

Dasari D, Hendriyanto A, Sahara S, Suryadi D,
Muhaimin LH, Chao T and Fitriana L (2024)
ChatGPT in didactical tetrahedron, does it
make an exception? A case study in
mathematics teaching and learning.
Front. Educ. 8:1295413.
doi: 10.3389/educ.2023.1295413

COPYRIGHT

© 2024 Dasari, Hendriyanto, Sahara, Suryadi,
Muhaimin, Chao and Fitriana. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

ChatGPT in didactical tetrahedron, does it make an exception? A case study in mathematics teaching and learning

Dadan Dasari^{1,2*}, Agus Hendriyanto^{1,2}, Sani Sahara¹,
Didi Suryadi^{1,2}, Lukman Hakim Muhaimin¹, Theodore Chao³ and
Laila Fitriana⁴

¹Department of Mathematics Education, Universitas Pendidikan Indonesia, Bandung, Indonesia,

²Indonesian DDR Development Center (PUSBANGDDRINDO), Universitas Pendidikan Indonesia,

Bandung, Indonesia, ³College of Education and Human Ecology, Ohio State University, Columbus,
OH, United States, ⁴Faculty of Teacher Training and Education, Universitas Sebelas Maret, Surakarta,
Indonesia

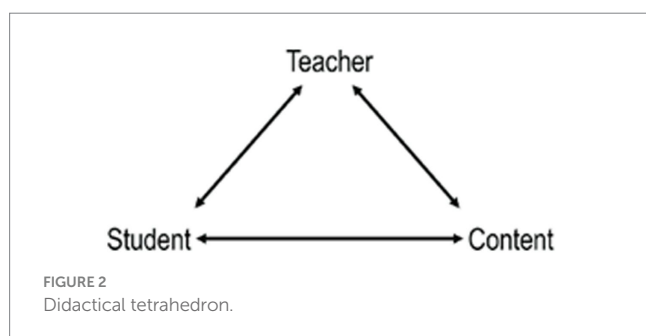
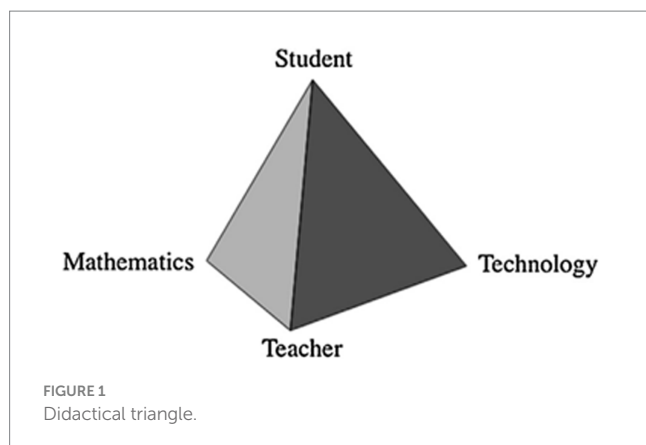
The didactical tetrahedron model proposes a framework for integrating technology into the previous didactical triangle. This study addresses this issue through examining the role of ChatGPT in educational settings. This quantitative and qualitative study reveals differences among three groups. We observed that students relying solely on ChatGPT for learning resulted in lower performance compared to those receiving instruction from teachers, either alone or supported by ChatGPT. The findings highlight the potential of ChatGPT in enhancing mathematical understanding, yet also underscore the indispensable role of instructors. While students generally perceive ChatGPT as a beneficial tool for learning mathematical concepts, there are concerns regarding over-reliance and the ethical implications of its use. The integration of ChatGPT into educational frameworks remains questionable within a didactic context, particularly due to its limitations in fostering deep information comprehension, stimulating critical thinking, and providing human-like guidance. The study advocates for a balanced approach, suggesting that ChatGPT can augment the learning process effectively when used in conjunction with guidance. Thus, positioning technology as an independent focal point in transforming the didactic triangle into a didactical tetrahedron is not appropriate, even when represented by ChatGPT.

KEYWORDS

ChatGPT, didactics, didactical tetrahedron, learning, teaching

1 Introduction

Previous research such as that conducted by [Freudenthal \(1991\)](#), [Brousseau \(1997\)](#), and [Steinbring \(2005\)](#), has focused on math learning and teaching by studying the interactions among students, teachers, and the subject matter. These interactions are typically represented using the didactic triangle (see [Figure 1](#)). The didactic triangle, also known as the “instructional triangle,” is an early model introduced by the German educator Friedrich Christoph Dahlmann



in the 1960s. It consists of three fundamental components: the teacher, the student, and the teaching material (Straesser, 2007). The model emphasizes the interconnectedness of these three elements in the process of teaching and learning. In the didactic triangle, the teacher is responsible for mediating and facilitating the transfer of knowledge to the students. Students play an active role in the learning process, engaging with the teaching material and interacting with the teacher. The teaching material represents the content and concepts being taught (Brousseau, 1997).

The didactic triangle provides a foundational framework for understanding the interaction among these three elements, underscoring the importance of a harmonious alignment for effective teaching and learning. To explicitly consider the role of technology in these interactions, the didactic triangle has been expanded by Tall (1986) and more recently by Olive et al. (2010) and Ruthven (2012) into a new concept known as the didactic tetrahedron (see Figure 2).

The didactical tetrahedron, a conceptual model used in education, particularly in the context of mathematics and digital technology integration, extends the traditional didactical triangle by incorporating technology as a critical fourth element (Jukić Matić and Glasnović Gracin, 2016; Prediger et al., 2019). This inclusion underscores the transformative role of technology in reshaping the educational landscape, especially in facilitating investigative approaches to teaching and learning mathematics (Remillard and Heck, 2014). It redefines the role of teachers from mere conveyors of knowledge to facilitators who guide students in a technologically enriched learning environment. The framework also brings to light the challenges of integrating technology in education, while simultaneously presenting opportunities for enhancing student engagement and understanding of mathematical concepts (Olive et al., 2010; Cao et al., 2021). Therefore, it offers a comprehensive framework for understanding the

complex interactions between teachers, students, subject matter, and technology in the modern educational context, advocating for a holistic approach that encompasses these interconnected dimensions to enrich the teaching and learning experience.

Since its introduction, the didactical tetrahedron has been widely embraced and expanded upon by researchers, educators, and curriculum developers. It has influenced the design of educational materials, the development of teaching strategies, and the integration of technology in various educational contexts (Aldon et al., 2021; Cao et al., 2021; Novita and Herman, 2021). This concept remains relevant and widely used in educational research and practice, offering a holistic perspective on the teaching and learning process. According to Olive et al. (2010), the introduction of technology into didactic situations can have transformative effects, leading to a better representation using the didactic tetrahedron. This tetrahedron illustrates the interaction between the teacher, students, tools, and mathematical knowledge, all mediated by technology (p. 168).

Suryadi (2019) provides criticism regarding the independence of technology as a new point in the didactic triangle. Through the question, “Can technology stand alone in relation to didactics?” Suryadi (2019) explains that we must return to the definition of didactics, where didactics is related to the diffusion and acquisition of knowledge. Because the actors of the diffusion process are teachers, the actors of the acquisition process are students, and the substance being diffused or acquired is existing knowledge (the result of transposition processes), the idea of the didactic triangle can logically be understood as the relationship between the three main entities of the diffusion and acquisition of knowledge events. In the didactic triangle, each party has its role in the context of diffusion and acquisition of knowledge, resulting in relationships that can be explained from both the diffusion and acquisition perspectives (Bosch, 2015).

The student-material relationship is referred to as the didactic relationship, which describes the process of knowledge acquisition in the didactic context as science, epistemology, or art (Suryadi, 2019b). The teacher-student relationship is called the pedagogical relationship because interaction between the educator and the learner is required in the process of diffusion and acquisition of knowledge (Suryadi, 2019b). This interaction is fundamentally based on the learner’s need for development, both in actual and potential stages, as explained in Vygotsky’s theory (Vygotsky, 1978). The teacher-material relationship is called the anticipatory didactic-pedagogical relationship (ADP). This relationship illustrates that an educator must have predictive thinking regarding the process of diffusion and acquisition of knowledge so that every possible outcome in the acquisition of new knowledge is anticipated, both from a didactic and pedagogical perspective.

Furthermore, Suryadi (2019) explains that technology plays a crucial role in the diffusion and acquisition of knowledge. However, it is important to realize that technology is not an independent entity but a tool used by educators to enhance the learning process (Ghavifekr and Rosdy, 2015). This requires thoughtful consideration and design by educators to make the educational experience more engaging, efficient, and easily accessible. The philosophical justification for the use of technology lies in its ability to facilitate extended cognition, providing opportunities for humans to develop their cognitive abilities. Meanwhile, in the didactical tetrahedron, the

interaction between technology and teacher, student, and content is described as follows:

1. Technology and teachers: Technology plays a crucial role in supporting teachers' teaching practices and enhancing their teaching methods. Teachers can use technology to access various digital resources, teaching aids, and multimedia materials that can enrich their lessons and make them more engaging.
2. Technology and students: Technology provides students with new opportunities to learn, collaborate, and express themselves. It allows students to access a vast amount of information, research materials, and educational content beyond traditional classroom resources. Through technology, students can engage in interactive and multimedia-rich learning experiences, which can enhance their understanding and retention of course materials. Furthermore, technology enables personalized learning experiences, adaptive assessment, and feedback, catering to individual student needs and promoting independent learning. Collaborative technology also facilitates peer-to-peer learning, communication, and teamwork.
3. Technology and content: Technology can transform and augment the materials used in teaching and learning (Alneyadi et al., 2023). It provides alternative formats, multimedia presentations, simulations, and interactive resources that make course materials more accessible and engaging. Digital textbooks, e-books, online databases, and educational websites offer extensive and up-to-date information on various topics. Additionally, technology enables the creation of digital learning materials such as educational videos, interactive presentations, and online quizzes tailored to specific learning objectives. Technology also allows real-time updates and modifications to materials, ensuring that they remain current and relevant (Olive et al., 2010; Rezat and Sträßer, 2012; Ruthven, 2012).

Overall, technology serves as a catalyst in the didactical tetrahedron, supporting teachers in their teaching practices, empowering students in their learning experiences, and enriching the materials used in the teaching process. When integrated thoughtfully and purposefully, technology can enhance educational outcomes and foster creativity, critical thinking, and collaboration between teachers and students (Yang and Wu, 2012). Thus, there is no urgency to place technology as an independent point in the didactic concept. But what about the recent technology we know as Chat-GPT? Does the presence of Chat-GPT justify that technology can stand independently in the didactic concept, making the didactical tetrahedron relevant?

ChatGPT is an advanced language model developed by Open AI (Biswas, 2023; Lund and Wang, 2023). It is built upon the GPT (Generative Pre-trained Transformer) architecture, specifically GPT-3.5 (Rehana et al., 2023). This model is designed to generate human-like text responses based on the input it receives in the form of prompts or messages (Haleem et al., 2022; Adiguzel et al., 2023; Pavlik, 2023). ChatGPT is trained on a large dataset of internet text, allowing it to learn language patterns, structure, and contextual understanding (Lund and Wang, 2023). It can comprehend and produce text across various domains, covering a wide range of topics,

with the goal of generating coherent and contextually relevant responses that simulate natural human conversation (Hassani and Silva, 2023). The chat-based GPT format allows users to engage in interactive and dynamic conversations with the model. Users provide instructions or messages, and the model generates responses based on the input it receives. The model's responses are not pre-determined but generated quickly, taking into consideration the conversation context. Open AI has provided various versions of the GPT model, and ChatGPT is one specific implementation focused on providing conversational capabilities (Mhlanga, 2023). It has been used in various applications, including customer support, language translation, creative writing assistance, and education support (Mattas, 2023).

In terms of educational assistance, ChatGPT, an advanced language model, has a significant role to play in the learning process. It assists in various ways, such as information retrieval, enabling learners to quickly access and expand their knowledge on a wide range of topics (Lo, 2023). When it comes to explaining concepts, ChatGPT excels by breaking down complex ideas into understandable components, providing examples, and offering clarifications, thus deepening learners' understanding (Coskun, 2023). As a practice partner, ChatGPT engages learners in simulated conversations or written exchanges, offering valuable feedback on grammar, vocabulary, and coherence, thereby enhancing communication skills (Shaikh et al., 2023). The interactive learning experiences facilitated by ChatGPT, such as quizzes, puzzles, and riddles, not only engage students but also allow for a more personalized learning journey through its adaptive responses (Elbanna and Armstrong, 2023).

Observing how technology is now perceived, especially given its rapid development over the past decade, it seems that ChatGPT, as a trained language model, has the potential to usher in a new reality about technology in the realm of education, and specifically in its position within this study's framework. Justifications about technology in education that were held previously now open up new discussions and questions as breakthroughs in technology emerge and evolve too quickly. These developments may lead to a fresh interpretation of technology compared to before. It is, therefore, very important for us to continually evaluate and explore its role in education. Such dynamics will help maintain the strength of knowledge and ensure that mathematics education remains epistemic for students (Gupta and Elby, 2011). This research aims to explore the potential use of Chat-GPT in mathematics education. Through a comprehensive study, the objective is to assess the role and impact of Chat-GPT on overall student performance, engagement, and learning experiences.

Therefore, we propose three research questions that we will answer through this study:

1. Is there a significant difference in math performance between students who solely use ChatGPT without any guidance from a lecturer, those who receive instructions with ChatGPT's assistance, and those who receive regular instructions without ChatGPT's help?
2. What do students think about using ChatGPT for learning and grasping mathematical concepts?
3. Does the inclusion of ChatGPT provide a valid justification within the didactical tetrahedron framework?

2 Materials and methods

Referring to the research question posed, this study adopts two approaches: quantitative research and qualitative research. We're not calling it a mixed-methods study because we believe that ontologically, quantitative and qualitative research are at odds with each other, making it challenging to combine them to investigate the same issue. Quantitative research is rooted in the ontological perspective known as positivism or postpositivism (Dieronitou, 2014). Positivism holds that there is an objective reality that can be studied and understood through empirical observation and measurement (Tuli, 2011). It argues that the social world operates according to generalizable laws, similar to those found in the natural sciences. In this view, reality is considered external and independent of the researcher, and its goal is to uncover universal patterns and cause-and-effect relationships. On the other hand, qualitative research is based on a different ontological perspective known as constructivism, interpretivism, or social constructivism (Lee, 2012). This perspective argues that reality is socially and subjectively constructed, and meaning is actively created by individuals and groups through their interactions and interpretations of the world (Fischer and Guzel, 2023). Qualitative research seeks to understand the complexity and depth of human experiences, perspectives, and social phenomena (Rahman, 2016). In short, in quantitative research, the relationship between the researcher and the research sample should be independent, whereas in qualitative research, the relationship between the researcher and the research subjects should be dependent.

In this case, the researchers conducted quantitative research first to address the first research question. During this phase, the researchers ensured that the relationship between the researcher and the research sample remained independent. Subsequently, the second and third research questions were answered using a qualitative approach.

2.1 Research design

In the quantitative research part, we adopted a static group comparison as part of a quasi-experimental design. This type of design is particularly useful when traditional experimental designs are not practical or ethical. In our case, it involved studying both an experimental group and a control group with different treatments, as described by Kirk (2009). The experimental group consisted of students who received instruction using ChatGPT. This study included two different experimental groups: experimental group 1, where students received complete instructions solely through ChatGPT during the learning process, and experimental group 2, where students engaged in collaborative learning with a teacher using ChatGPT as a tool. Meanwhile, the control group consisted of students who received treatment as usual. In the quasi-experimental design, we compared groups or conditions that already existed (e.g., different classes or schools), without randomly assigning participants to conditions or manipulating the independent variable. This approach was chosen due to the logistical constraints within schools and educational systems that often make traditional experimental designs difficult to implement. In practice, we enlisted local teachers to administer the treatments, ensuring that the treatment for the control group closely resembled their regular classroom experience. However, prior to the

study, researchers also provided guidance to the respective teachers regarding the treatment for experimental group 2, which involved the integration of ChatGPT into the learning process.

On the other hand, in the qualitative research part, we employed a phenomenological design. Phenomenological qualitative research is an approach aimed at understanding the life experiences and subjective perspectives of individuals (Creswell, 2012). It seeks to explore and describe the essence and meaning of the experiences that students go through when using ChatGPT in their mathematics learning. Phenomenology focuses on phenomena that emerge in consciousness and emphasizes understanding the rich and unique qualities of an experience.

2.2 Sampling and subjects

The quantitative research was conducted in the city of Surakarta, Indonesia, involving a study population comprising students from three universities, each with the same accreditation level. These participants were sixth-semester mathematics education students enrolled in a numerical methods course. Additionally, they had undergone a relatively similar selection process for admission, in terms of both content and procedure. Due to logistical limitations and the inability to assign participants randomly, a cluster random sampling was used. One of the universities was selected as the sample, with just one class currently undertaking the numerical methods course. In the next phase, the chosen class for the study was divided into three groups through random sampling of 33 papers. Each paper contained information about its group type, and there were 11 papers for each group type in total. This random sampling of the 33 papers created a more detailed experimental framework. Class A, consisting of 9 participants, was designated as experimental group 1. Similarly, Class B, also with 9 participants, was identified as experimental group 2. Meanwhile, Class C, with 11 participants, was set as the control group. Thus, this study involved 29 undergraduate students in a mathematics education program.

Meanwhile, the qualitative research involved in-depth interviews with a total of five students who had recently used ChatGPT in their learning process. The participants were selected using a combination of criteria and snowball sampling techniques. Inclusion criteria for participants included being part of the experimental classes that had implemented ChatGPT in their learning and their ability to articulate their experiences in a thoughtful and reflective manner. To protect the confidentiality and anonymity of the participants, pseudonyms were assigned to each participant in the reporting of findings. Participants were provided with detailed information about the research, its purpose, and the voluntary nature of their participation. Informed consent was obtained from each participant before the interviews, ensuring their understanding of the research's objectives and their rights as research subjects (Marshall et al., 2006).

2.3 Data collection and instrument

The data for this research was collected using a combination of quantitative and qualitative methods tailored to the research questions. The aim was to provide a comprehensive understanding of the impact of the new instructional intervention (ChatGPT) on students'

mathematical problem-solving abilities, their attitudes toward using ChatGPT for learning mathematics, and ultimately provide justification for the role of ChatGPT in the didactic concept.

Quantitative data collection: To assess students' mathematical problem-solving abilities, we employed a posttest-only, non-equivalent control group design. This is a type of quasi-experimental research design that involves comparing the outcomes or effects of an intervention or treatment between two groups: the experimental group and the control group (Kirk, 2008). In this design, participants in all three groups were measured on the dependent variable (desired outcomes) after the intervention was administered, but there was no pretest measurement. Posttest scores were analyzed using statistical methods to test for significant changes and differences in students' math performance.

Qualitative data collection: Qualitative data were gathered to gain insights into students' attitudes toward ChatGPT and their experiences with the intervention. Semi-structured interviews were conducted with some participants to explore their perceptions, beliefs, and experiences related to learning mathematics and the instructional intervention. Interviews were audio-recorded and transcribed verbatim for analysis. The interview questions focused on students' attitudes toward ChatGPT, their involvement in the intervention activities, and their perceptions of the intervention's impact on their learning experience. Additionally, classroom observations were conducted to provide contextual information about the implementation of the intervention. Researchers observed the learning sessions to gather data on teaching methods, materials used, and student interactions. Field notes were taken during observations, capturing class dynamics, student engagement, and any significant observations related to the implementation of the intervention.

2.4 Data analysis technique

Quantitative Analysis: To assess the impact of the intervention on students' math performance, the posttest scores from the experimental and control groups are the focus of statistical analysis. First, descriptive statistics such as mean, standard deviation, and frequency are calculated to summarize the data and provide an overview of the students' math performance. Then, inferential statistical tests are used to determine whether there are any significant differences in math performance among the three groups. Specifically, a One-way ANOVA test is conducted to compare the average scores between the experimental and control groups (Kirk, 2008). Additionally, effect sizes are calculated to assess the practical significance of any observed differences.

Qualitative Analysis: Qualitative data collected through interviews and classroom observations are analyzed using thematic analysis. Transcripts of interviews and field notes from observations are carefully reviewed and coded to identify recurring themes and patterns related to students' attitudes toward ChatGPT and their experiences with instructional interventions. The coding process involves assigning meaningful labels to segments of data, grouping similar codes into categories, and refining the coding scheme through an iterative process. The identified themes are then interpreted and supported by relevant quotes from interviews and field notes to provide a nuanced and rich understanding of the students' perspectives.

3 Result and discussion

In this section we briefly describe the result of three different groups in learning activities, then continue with the discussions regarding the mathematical performance, students' perspective about ChatGPT for learning and constructing mathematical concept, and the view of ChatGPT in didactics concept.

The first group that learning activities of the group using only ChatGPT as their primary instruction tool started off with enthusiasm. This method provided quick feedback on questions related to the numerical methods course. Unfortunately, this initial enthusiasm did not last long because the group became confused about what they should be asking the system. As a result, when they faced a test without ChatGPT's assistance, they were not prepared to tackle the challenges in problem-solving.

Now the group that used ChatGPT as a learning aid with the guidance of a facilitator, the lecturer, showed more directed results. The lecturer provided guidelines on the material to be studied, enabling students to utilize ChatGPT more effectively. Interacting with the lecturer allowed students to validate the information obtained from ChatGPT and receive additional explanations if there was any confusion. Thus, the knowledge acquired from ChatGPT could be directly validated by the lecturer for accuracy. Furthermore, this approach provided effective learning where the lecturer remained actively involved in the learning process while leveraging technology.

On the other hand, in groups that solely relied on the lecturer for learning, a more traditional approach was apparent. Nevertheless, this method resulted in a better understanding of the material compared to a method without lecturer involvement. Direct interaction with the lecturer allowed students to clarify doubts and, therefore, gain a deeper understanding. While the feedback may not be as quick as what ChatGPT offers, the closeness of interaction provided by the lecturer remained invaluable. This underscores that even though technology plays a role in enhancing the learning process, communication remains a key element in grasping the material.

3.1 Students' mathematical performance

This section aims to compare the math performance among students who exclusively used ChatGPT, students who received instructions from a lecturer with the assistance of ChatGPT, and students who only received instructions from a lecturer. The math instructions given to the students with the instructors were identical in terms of content, duration, and difficulty level. After the instructional period, an evaluation test consisting of a series of math questions covering the taught material was conducted. Test scores represented students' math performance. Table 1 outlines the details of the descriptive analysis results.

TABLE 1 Descriptive data.

Factor	N	Mean	SD	SE	Coefficient of variation
Experiment 1	9	9.816	5.553	1.851	0.566
Experiment 2	9	42.593	18.166	6.055	0.426
Control	11	30.303	13.288	4.006	0.438

Based on Table 1, it is observed that Experimental Group 1, which used only ChatGPT, achieved an average score of 9.816 with a standard deviation of 5.553. Experimental Group 2, which received instruction from a professor aided by ChatGPT, attained a much higher average test score of 42.593 with a standard deviation of 18.166. In contrast, the control group, receiving only regular instruction, had an average score of 30.303 with a standard deviation of 13.5288.

Prior to hypothesis testing, assumption tests were conducted, including tests for normality and homogeneity. Referring to Figure 3, the Q-Q Plot indicates that the data appears to be normally and linearly distributed. However, in the homogeneity test, the value of p was found to be 0.006, which is less than the alpha level of 0.05, indicating that the data is heterogeneous. Therefore, while the assumption of normality is met, the assumption of homogeneity is not. Consequently, the Welch ANOVA test was chosen to compare the means between groups in the experimental design, with the following hypothesis:

$$H_0 : \mu_A = \mu_B = \mu_C$$

$$H_1 : \text{except } H_0$$

The output of the testing using SPSS is presented in Table 2. According to Table 2, it is evident that the value of p is 0.001, which is less than the alpha level of 0.05, leading to the rejection of the null hypothesis (H_0). This indicates that there is a significant difference in the mean scores among the three groups. Furthermore, the Effect Size (ω^2) is calculated to be 0.469, suggesting a substantial impact (referenced). Due to the rejection of H_0 , *Post-hoc* Tests were conducted to identify which pairs of group means differ significantly (see Table 3).

The *post-hoc* tests reveal significant differences between Class A and Class B (p_{tukey} value = $0.001 < \alpha = 0.05$), and between Class A and Class C (p_{tukey} value = $0.006 < \alpha = 0.05$). However, no significant difference was found between Class B and Class C (p_{tukey} value = $0.122 > \alpha = 0.05$). However, given the difference in means, the use

of ChatGPT as a teaching aid in mathematics education supplemented by instructor guidance, is worthy of further investigation.

Quantitatively, it is evident that the math performance of students taught by lecturers or with ChatGPT assistance is better compared to students who learn without a lecturer and solely rely on ChatGPT. This underscores the importance of the lecturer's role in providing deeper and contextual math instruction, while ChatGPT can offer additional support in understanding concepts. This research outcome indicates the significance of the lecturer's role in delivering context-rich, personalized instruction with direct interaction. Lecturers can adapt their teaching approach to students' needs, clarify complex concepts, and stimulate in-depth discussions. Therefore, lecturer-student interaction remains a significant factor in enhancing students' understanding of mathematics, as conveyed by Heggart and Yoo (2018).

Teachers (lecturers) possess a domain of knowledge and expertise in the field of mathematics, enabling them to provide tailored guidance, explanations, and clarifications to meet students' needs (Troussas et al., 2020). ChatGPT can offer information and answers based on its data, but it may not have the same level of expertise or understanding of individual student needs as lecturers do (Baidoo-Anu and Owusu Ansah, 2023; Lecler et al., 2023). ChatGPT also lacks the ability to provide justification regarding whether the information it conveys holds absolute truth (Sun and Hoelscher, 2023). Additionally, lecturers can adapt their teaching methods based on students' progress and learning styles, providing direct feedback. They can offer personalized instructions and modify their approaches to cater to different learning needs. On the other hand, ChatGPT can provide standard responses without the ability to adapt to each student's learning requirements (Aydin and Karaarslan, 2023).

Mathematics can be a complex subject, and students often require in-depth explanations and clarifications of abstract mathematical concepts. Lecturers can provide real-time examples, demonstrations, and interactive discussions to help students grasp these mathematical ideas. While ChatGPT can provide information, it cannot offer the same level of dynamic and interactive explanations (Baidoo-Anu and Owusu Ansah, 2023; Ray, 2023). Lecturers are trained in pedagogical strategies and teaching methodologies designed to enhance student learning (Phuong et al., 2018). They can employ various teaching techniques, such as visual aids, problem-solving exercises, and interactive activities, to engage students and foster a deeper understanding (Brinkley-Etzkorn, 2018; Singh et al., 2021). ChatGPT, as a language model, lacks the same pedagogical training and cannot effectively use these strategies (Kasneci et al., 2023).

Low performance in groups solely relying on ChatGPT indicates limitations in the model's ability to provide comprehensive math learning support. ChatGPT may struggle with understanding highly specific questions, offering context-appropriate examples, and solving more complex problems (Tlili et al., 2023). Study by Hassan et al. (2023) reported that ChatGPT can handle routine inquiries and tasks, even though it is only freeing up time for more complex task. Also, Ray (2023) discussed on how handling more complex tasks may still present challenges. It is essential to note that while ChatGPT can provide valuable information and assistance, it is most effective when used as a complement to human instruction rather than a substitute for human educators (Jeon and Lee, 2023; Tlili et al., 2023). The combination of human expertise and guidance with ChatGPT's capabilities presents an opportunity to enhance students' math

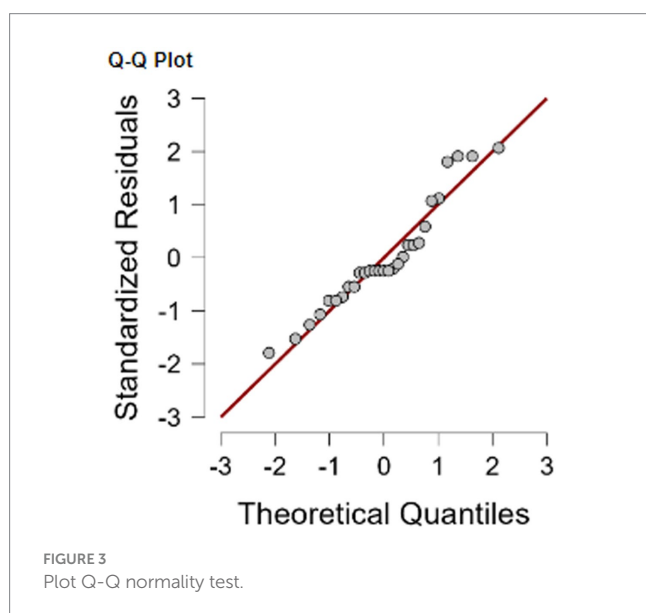


TABLE 2 ANOVA test.

Homogeneity correction	Cases	Sum of squares	df	Mean square	F	p	ω^2
Weich	Faktor	4,949,398	2,000	2,474,699	20,600	< 0,01	0,469
	Residuals	4,652,302	14,441	322,155			

Type III sum of squares.

TABLE 3 *Post-hoc* comparisons.

		95% CI for mean difference					95% CI for Cohen's d			
		Mean difference	Lower	Upper	SE	t	Cohen's d	Lower	Upper	P _{tuke}
Control	Experiment 2	20.487	5.547	35.427	6.012	3.408	1.532	0.259	2.804	0.006
	Experiment 2	-12.291	-27.231	2.649	6.012	-2.044	-0.919	-2.114	0.277	0.122
Experiment 1	Experiment 2	-32.778	-48.447	-17.109	6.306	-5.198	-2.450	-3.937	-0.963	<.001

performance. This finding highlights the potential use of ChatGPT technology as an effective learning tool in improving students' math performance. Further research can explore optimal ways to integrate this technology into broader learning contexts.

3.2 Students' perspective about ChatGPT for learning and constructing mathematical concept

This section aims to investigate students' views on the effectiveness and benefits of using ChatGPT in the learning process and understanding mathematical concepts. The interviews focus on questions related to their experiences using ChatGPT in the context of learning mathematics, their opinions on the utility of the tool, and the impact of using ChatGPT on their understanding of mathematical concepts. There are 15 main questions posed to explore students' perspectives and experiences regarding the use of ChatGPT in the learning process and the formation of their understanding of mathematical concepts. Table 4 presents the questions asked to the students.

The thematic analysis of the interviews reveals some important findings. Overall, the students were familiar with ChatGPT before this research was conducted. However, some students had limited knowledge about ChatGPT's functions. In general, they described this technology as a computer program that uses artificial intelligence (AI) to generate text or respond to user requests. As a result, most of them used ChatGPT mainly to complete tasks related to paper writing or essays. Using ChatGPT to assist in essay writing could be a useful tool for students, especially when they have difficulty in constructing sentences or expressing ideas effectively.

Meanwhile, some students have used ChatGPT as a tool to understand mathematical concepts. Students who were taught by instructors with the help of ChatGPT found it to be a useful tool in learning mathematics. They appreciated ChatGPT's ability to provide clear and structured explanations of complex mathematical concepts. Students also highlighted ChatGPT's ability to provide solutions and problem-solving strategies that were useful in their math assignments. Students reported that using ChatGPT improved their understanding of mathematical concepts. They felt more confident and capable of

overcoming difficulties in understanding with the help of ChatGPT. They also emphasized that ChatGPT's use is suitable in the context of providing explanations about basic concepts, definitions, and procedures. This aligns with ChatGPT's main strength in generating text that can break down concepts in a structured way.

Some students felt that ChatGPT helped them deepen their understanding of the material they were studying. They considered it a valuable resource for further understanding challenging aspects of the concepts they found difficult. However, besides these benefits, some students also expressed concerns about excessive reliance on ChatGPT. They argued that excessive dependence on technology could reduce their ability to solve problems independently. Additionally, it should be noted that ChatGPT's effectiveness is currently limited in understanding complex mathematical contexts or providing highly personalized explanations. Moreover, while ChatGPT serves as a helpful tool, its limitations in comprehensibility and adaptability to individual learning styles cannot be overlooked. The information provided by ChatGPT, though extensive and varied, may not always align perfectly with the specific curricular context or the unique conceptual misunderstandings a student might have. This gap can lead to partial or misaligned understanding, especially in subjects where nuance and depth of knowledge are critical. Furthermore, ChatGPT's algorithmic nature means it might not always capture the subtleties of human thought processes or the specific pedagogical approaches that a teacher might use to address a student's unique learning needs. This limitation underscores the importance of using ChatGPT as a supplementary tool, one that complements but does not replace the personalized guidance and expertise of a human educator.

In some cases of complex math problems, students revealed their inability to understand ChatGPT's responses. The explanations provided by the system were not consistent with the students' prior knowledge, coupled with ChatGPT's inability to solve more complex math problems. Consequently, students who should have benefited from ChatGPT ended up facing obstacles in their learning. Hence, there is a need for instructors to play a role as knowledge confirmers. However, for simpler math cases, students acknowledged the help provided by this system, allowing them to learn more independently and acquire various ways or tricks to answer questions. Some students added that the clarity and

TABLE 4 Questions to understand students' perspectives and experiences related to the use of ChatGPT.

Question	Aspect
Have you ever used or interacted with ChatGPT before? If yes, in what context?	Introduction and Use of ChatGPT
How would you describe ChatGPT to someone who has never heard of it? What's your initial understanding of this technology?	
Have you ever used ChatGPT as a tool to understand mathematical concepts? If yes, how was your experience?	Learning Math with ChatGPT
Are there specific situations where you feel ChatGPT could help you understand difficult math material?	
What kind of math questions or problems do you think are suitable for solving with ChatGPT? Why?	
In your opinion, what are the main advantages of using ChatGPT in math learning compared to traditional methods?	Advantages and Disadvantages of ChatGPT
Do you identify any drawbacks or limitations in using ChatGPT to understand math concepts? If yes, what are they?	
How do you interact with ChatGPT when trying to explain or understand complex math concepts?	Interaction and Personal Experience
Could you provide a concrete example of how ChatGPT has helped you overcome a barrier in understanding a math concept?	
To what extent do you find ChatGPT effective in helping you understand math concepts deeply? Have there been any changes in how you understand and apply these concepts after interacting with ChatGPT?	Effectiveness and Learning Outcomes
Are you concerned that relying too much on ChatGPT might reduce your ability to understand math concepts independently? Why or why not?	Ethical Considerations and Dependency
What are your thoughts on the importance of developing critical thinking skills and problem-solving abilities in math learning, regardless of tools like ChatGPT?	
Do you have any suggestions for further development of ChatGPT to make it more effective in aiding math learning and understanding?	Suggestions and Improvements
Are there any specific features or functionalities you would expect from ChatGPT to make it more suitable for use in a math learning context?	
How do you compare the use of ChatGPT to other math learning resources like textbooks, instructional videos, or instructors?	Comparison with Other Alternatives

accuracy of ChatGPT's answers depend on the clarity and detail of the questions asked. The clearer and more detailed the questions, the closer the response will be to what they are looking for. However, sometimes students accept ChatGPT's concepts without any justification from the instructor or other sources. Therefore, learning would be more effective if ChatGPT is used as a tool to assist learning under the guidance of instructors, as ChatGPT cannot provide a human-like guidance. This reliance on ChatGPT for learning math without proper guidance or confirmation from educators can lead to misunderstandings and incomplete learning. Students, particularly those with less experience in self-directed learning or weaker foundational knowledge.

Next, there are students' concerns about ethical considerations. They feel that while AI technology offers benefits, it also has the potential to raise ethical dilemmas. These students worry that excessive reliance on ChatGPT could hinder the development of their skills. There is a common sentiment that the lure of convenience may come at the expense of intellectual growth and contribute to the spread of biased information. Concerns about addiction are also felt by some students, especially after observing examples where they and their peers became overly dependent on ChatGPT for tasks that could be done independently. On the contrary, other students do not share these concerns and consider AI tools as valuable assets that greatly assist their academic achievements. They appreciate its ability to generate ideas, answers, and explanations quickly, allowing them to delve deeper into complex subjects.

The results show a dual perspective among students regarding ChatGPT, making it clear that while it offers undeniable benefits, there are growing concerns about its ethical dilemmas. This aligns with

research conducted by [Geis et al. \(2019\)](#), [Gong et al. \(2019\)](#), [Pedro \(2019\)](#), [Safdar et al. \(2020\)](#), and [Sit et al. \(2020\)](#), which surveyed students' attitudes toward AI-assisted learning tools, especially in terms of ethics and its implications, including privacy, bias, and transparency issues. This is consistent with the findings of this study, where students expressed their concerns about the potential for AI to contribute to the spread of biased information. Additionally, students also voiced concerns that excessive reliance on ChatGPT could hinder their intellectual development. This sentiment experienced by students aligns with [Alam \(2022\)](#) research on the impact of AI tools, showing that students who rely too heavily on AI-generated content demonstrate a decline in their ability to independently analyze and synthesize information. This indicates that while AI tools can provide quick solutions, they may impede the development of their cognitive skills ([Vincent-Lancrin and van der Vlies, 2020](#)).

As AI technology continues to evolve and integrate into various aspects of society, there is an urgent need to address the ethical issues it raises. Institutions should consider strategies to promote balanced AI tool usage while also encouraging ethical considerations. However, some argue that AI is the future ([Gautam et al., 2022](#)), and schools should embrace it rather than restrict it. At least, students view ChatGPT as certainly not the end of the world and not a complete suppression of independent thinking.

Overall, this section demonstrates that students have a positive perception of the usefulness and effectiveness of ChatGPT in learning and understanding mathematical concepts. Despite some concerns, the use of ChatGPT is seen as a valuable tool in enhancing mathematics education. This research provides valuable insights for the development of ChatGPT technology in an educational context

and underscores the importance of considering its limitations and appropriate utilization.

3.3 ChatGPT in didactics concept

The two subtopics discussed above have highlighted the immense potential of ChatGPT to engage in the student learning process, especially in mathematics education. This section will further elaborate on how ChatGPT's potential is viewed from a didactic perspective. In the didactic triangle, it has been explained that the connecting line between the teacher and the student is called the pedagogical relationship, the line between the student and the material is called the didactic relationship, and the line between the teacher and the material is called the didactic pedagogical anticipation (Suryadi, 2019a). By adding a new independent point, which is technology, to the didactical tetrahedron, new areas emerge: (1) teacher-student-technology; (2) student-material-technology; and (3) teacher-material-technology. Additionally, new lines emerge as well: (1) teacher-technology; (2) student-technology; and (3) material-technology. However, in its justification (Ruthven, 2009; Olive et al., 2010), both in the field and lines of technology, it is only seen as a tool to optimize the learning process. There is no comprehensive explanation as in the case of the fields and lines in the didactic triangle (teacher-student-material), so technology cannot be considered an independent point like the others.

Specifically, this section discusses the inclusion of ChatGPT (as a form of technology) in the didactical tetrahedron as an independent point. Why do not we hold the same opinion as with other educational technologies? This is because of how ChatGPT operates, which is capable of answering various questions posed to it. This discussion involves findings from the previous section. Referring to quantitative data, it is known that a group of students who only use ChatGPT without any intervention from a teacher can also generate knowledge, enabling them to achieve scores not significantly worse than two other groups. The question is, can using ChatGPT alone really build their knowledge?

A group of students who only use ChatGPT, without any intervention from a teacher, may still be able to generate knowledge. However, there are some considerations to keep in mind. Even though ChatGPT can provide useful information and stimulate understanding, the knowledge obtained may not be complete and systematically organized. This limitation primarily occurs because ChatGPT does not provide comprehensive information due to limitations in constructing sentences and other constraints (Ray, 2023). Therefore, the ability to develop cohesive and in-depth knowledge may be hindered if relying solely on ChatGPT (Jarrah, et al., 2023). Furthermore, it is important to remember that building substantial and structured knowledge in a discipline involves more than just receiving information (Langer, 2011). Students need guidance and direction from experts (such as teachers) who can help them understand complex concepts, address misconceptions, and steer them toward deeper resources (To and Carless, 2016). Teachers can also help ensure that students understand the limitations of studying a field and provide the necessary context for better understanding (Yu, 2023). The material presented by ChatGPT may only be a starting point that needs to be further explored through discussion, analysis, and deeper exploration.

Before delving further into this phenomenon, it is worth recalling the process experienced by Plato when learning from Socrates, where Plato's learning with Socrates can be characterized by the "Socratic method" of questioning (dialectic). It is a method where Socrates engaged others in dialog by asking probing questions to stimulate critical thinking and examine one's beliefs and assumptions (Owens, 1950). This method is similar to what can be achieved through interaction between students and ChatGPT. Any "question" can be answered by the AI's reliability, so when the questioner has a critical mindset, it is very possible that ChatGPT's role can build new knowledge for them. Thus, can we conclude that the presence of ChatGPT can provide justification within the didactical tetrahedron?

It is important to note that the mathematics performance in this research is derived from the students' work on essay questions. Assessment of essay questions is subjective, meaning it is not just a binary right or wrong but also evaluates the process. There is no guarantee that the "score" obtained by students is solely the result of knowledge gained through their interactions with ChatGPT. Therefore, further analysis is needed. The first analysis is related to how the knowledge formed by each student on the same material. This is done by analyzing the questions posed by students and the answers provided by ChatGPT. Here's an explanation.

The chosen material is about solving ordinary differential equations (ODE) with the Runge-Kutta and Euler methods. Some questions asked by students and the answers provided by ChatGPT are presented in Table 5.

Table 5 presents some findings regarding ChatGPT's responses to questions posed by various subjects about the Euler method. Even though these questions revolve around the same concept, which is the Euler method, ChatGPT's responses vary in terms of how the material is presented and explained. When analyzing these responses, it can be observed that ChatGPT tends to provide summaries of the Euler method. The knowledge presented in these responses is not presented systematically or epistemically. This has the potential to create obstacles in the learning process, and it may even lead to misconceptions in understanding the material. For example, the explanation about the Euler method equation displayed is: "next step solution = previous step solution + (gradient \times step size)" (K132001), $y_{i+1} = y_i + h \times f(x_i, y_i)$ (K132002), $y(i+1) = y(i) + h \times f(x(i), y(i))$ (K132003), and $y(x+h) \approx y(x) + h \times f(x, y)$ (K132004).

In case K132001, there is a difference in the presentation model used by ChatGPT. This response uses a geometric interpretation to explain the Euler method equation. However, without additional explanation, this interpretation might be challenging for students to grasp without further insight into the origins and the fundamental ideas behind the equation. Other cases, such as K132002, K132003, and K132004, also indicate difficulties in understanding that the equations generated are derivatives of the expression $y(x_{i+1})$ with respect to x_i using the Taylor series. Explanations may be necessary to help students connect the equations with the underlying basic concepts.

In this context, ChatGPT's ability to provide adequate explanations depends on the suitability and clarity of the questions asked by the user. If these questions are asked more precisely and appropriately, ChatGPT is likely to provide more comprehensive and structured explanations about the Euler method concept and how the equations emerge from deeper reasoning. It is important to highlight that the

TABLE 5 Questions to understand students' perspectives and experiences related to the use of ChatGPT.

Orders or questions	ChatGPT answers	Subject ID
Euler's numerical method	<p>The Euler Numerical Method is an iterative approach used to solve ordinary differential equations (ODEs). It's named after the Swiss mathematician Leonhard Euler, who developed it in the 18th century. The basic idea of this method is to divide a specific time interval into small steps and approximate the solution's value at each of these steps.</p> <p>The main steps in the Euler method are as follows:</p> <ol style="list-style-type: none"> 1 Determine the ordinary differential equation to be solved and the initial condition (the solution's value at the starting point). 2 Choose the time interval to be used and divide it into small steps using a step size (h). 3 Use the differential equation to calculate the gradient or rate of change at the initial point. 4 Calculate the solution's value at the next step using the equation: solution at the next step = solution at the previous step + (gradient \times step size). 5 Repeat steps 3 and 4 for each subsequent step until reaching the desired endpoint. <p>The Euler method has a relatively high error rate and is usually used as an initial method or as a basis for more sophisticated numerical methods. It can provide adequate results for simple differential equations or when a very small step size is used.</p>	K132001
Euler's Method	<p>The Euler method is a numerical technique used to approximate the numerical solution of ordinary differential equations (ODEs). The basic principle of the Euler method involves approximating the first derivative of a function using a linear approach around a known starting point. In the Euler method, the differential equation is expressed as a first-order ordinary differential equation:</p> $\frac{dy}{dx} = f(x, y)$ <p>with the initial condition $y(x_0) = y_0$. To use the Euler method, the steps are as follows:</p> <ol style="list-style-type: none"> 1 Determine the range of x values for which you want to find the solution (e.g., from x_0 to x_n). 2 Specify the initial value y_0 for the initial condition $y(x_0)$. 3 Choose a time step or step size h (e.g., 0.1 or 0.01) to be used in the iteration process. 4 Start with the initial values x_0 and y_0. 5 Use the differential equation $f(x, y)$ to calculate the value of the derivative $\frac{dy}{dx}$ at the point (x, y). <ul style="list-style-type: none"> • Use the calculated $\frac{dy}{dx}$ value to update the value of y using the iteration formula: $y_{i+1} = y_i + h \times f(x_i, y_i)$, where i is the current iteration index, h is the time step, and x_i and y_i are the x and y values from the previous iteration. <ol style="list-style-type: none"> 6 Continue repeating steps 5 and 6, updating x with the time step h, i.e., $x_{i+1} = x_i + h$, until you reach the specified x_n. <p>The Euler method provides a simple and easy-to-implement numerical solution for ordinary differential equations, but it typically yields lower accuracy compared to more advanced methods like the Runge-Kutta method.</p>	K132002
Euler's Method	<p>The Euler method, also known as the tangent method, is one of the numerical methods used to solve ordinary differential equations (ODEs) numerically. It is named after the mathematician Leonhard Euler, who developed it. The Euler method is used to approximate the numerical solution of first-order ordinary differential equations with given initial conditions. Ordinary differential equations are typically expressed as:</p> $\frac{dy}{dx} = f(x, y)$ <p>Here, y is the function we want to find, x is the independent variable, and $f(x, y)$ is the function that defines the relationship between x, y, and the first derivative of y with respect to x.</p> <p>Here are the general steps in the Euler method:</p> <ol style="list-style-type: none"> 1 Determine the interval of x to be evaluated and the number of steps (h) to be taken within that interval. 2 Specify the initial value y_0 at the starting point x_0. 3 Calculate the value of y at each step using the following iterative formula: $y(i+1) = y(i) + h \times f(x(i), y(i))$, where $y(i)$ is the estimated value of y at step i, $x(i)$ is the x-coordinate at step i, and $f(x(i), y(i))$ is the value of the function f at step i. 4 Repeat step 3 for the desired number of steps. <p>The Euler method provides a simple linear approximation to compute the value of y at each step by utilizing the first derivative of y with respect to x and assuming that higher-order derivatives are not significant. Therefore, this method tends to yield accurate results only for relatively small intervals or when $f(x, y)$ is a linear or approximately linear function within the given interval.</p>	K132003

(Continued)

TABLE 5 (Continued)

Orders or questions	ChatGPT answers	Subject ID
Principles of the Euler method	<p>To understand the Euler method, we first need to comprehend the basic concept of differential equations. Differential equations are mathematical equations involving the derivatives of an unknown function. Ordinary Differential Equations (ODEs) are differential equations that involve only one independent variable. The Euler method employs an iterative approach to obtain a numerical solution for a differential equation. The steps of the Euler method are as follows:</p> <ol style="list-style-type: none">1 Determine the differential equation you want to solve. For example, let us consider an ODE: $y'(x) = f(x,y)$, where y is the function we seek, and f is the known function.2 Determined the initial value, $y(x_0) = y_0$, where x_0 is the starting point, and y_0 is the initial value of the function y at that point.3 Choose a step size h, which determines how small the increment in x will be in the iterative approximation. Smaller h values lead to more accurate Euler approximations.4 Use the ODE to derive an approximate differential equation: $y(x+h) \approx y(x) + h \times f(x,y)$.5 Apply the approximate differential equation at the given point (x,y). In each iteration, you can compute the value of $y(x+h)$ using the known $y(x)$ and its function.6 Repeat step 5 for each iteration until you reach the desired endpoint. <p>The Euler method is a simple and easy-to-understand numerical method. However, it's important to note that this method can yield inaccurate results for certain types of differential equations, especially if the value of h is too large or if there are rapid changes in the desired function. Therefore, there are more accurate numerical methods like the Runge-Kutta method that are also used for solving differential equations.</p>	K132004

Commands or questions and answers were delivered in Indonesian language.

information generated by ChatGPT comes from texts found on the internet (Javaid et al., 2023). Therefore, ChatGPT's responses are based on language patterns, information, and viewpoints present in these texts (Lund and Wang, 2023). The internet resources used by ChatGPT are vast and diverse, covering a wide range of texts, including news, articles, encyclopedias, blogs, forums, websites, and more (Apostolopoulos et al., 2023). In the learning process, this model attempts to understand the relationships between words, phrases, and ideas that appear in these texts. However, it is essential to remember that although ChatGPT can generate coherent and relevant texts based on what it has learned, it does not possess understanding in the conceptual and epistemological sense (Rees, 2022; Mitrović et al., 2023). Therefore, ChatGPT cannot definitively distinguish between valid and false theories.

This occurs because GPT is based on statistics and patterns in the training data and lacks the ability to make contextual judgments and deep understanding of the truth of a statement (Zhang et al., 2023). While ChatGPT can produce responses that sound logical and reasonable based on its learned data, it does not guarantee that these responses are always correct or in line with scientific truth (Wittmann, 2023). When interpreting information provided by ChatGPT, it is important to always consider other sources of information, conduct further research, and use personal judgment to determine the validity of a theory. ChatGPT can be a useful tool for generating text and potential information, but the ultimate responsibility for evaluating the truth and reliability of information remains with the user (Javaid et al., 2023). Therefore, the interaction between students and ChatGPT is significantly different from the Socratic method.

The learning process in the Socratic method involves (Morrison, 2010): (1) Wonder, where questions are posed. Socrates would initiate conversations or dialogs with Plato and other participants, approaching individuals to discuss various topics such as ethics, politics, and knowledge. Similar to ChatGPT, this system will work best if users start by asking questions (Tlili et al., 2023); (2)

Hypothesis, which is the response to wonder, where someone provides an opinion or statement regarding a question that becomes the hypothesis of the dialog. This differs from acquiring knowledge through ChatGPT, as the system lacks pedagogical capabilities similar to educators (Wardat et al., 2023), so the responses given by ChatGPT depend on what users ask; (3) Elenchus, refutation, and cross-examination. The essence of Socrates' practice is that hypotheses are questioned, and counterexamples are provided to prove or disprove the hypotheses. While ChatGPT is a text-based system based on user questions, its responses come from a wide range of internet sources, often with unclear origins (sources not cited), and users often accept the information provided by ChatGPT without cross-checking from other sources (Oviedo-Trespalcacios et al., 2023); (4) Acceptance or rejection of the hypothesis, where the choice is made between accepting or rejecting the counterexample information. Socratic methods emphasize sharp questioning and argument testing to gain a deeper understanding of concepts and sift valid from invalid information. On the other hand, interactions with ChatGPT often involve the consumption of raw information without much critical consideration. This can be a challenge if the information provided by ChatGPT is not entirely accurate, complete, or contextually suitable (Eriksson and Larsson, 2023); (5) Action, where the findings of the investigation are acted upon. In the Socratic method, taking action on findings is important for gaining a deeper understanding of arguments, identifying argument foundations, and exploring implications of discussed ideas. Socrates would encourage his interlocutors to think more critically and reflect on the arguments presented, allowing for a better understanding of the strengths and weaknesses of those arguments. On the other hand, in interactions with ChatGPT, users like students often tend to accept the information provided without much critical thought (Deiana et al., 2023). This can be a challenge if the information provided by ChatGPT is not entirely accurate, complete, or suitable within a specific context. Therefore, it is important for users to

develop critical thinking skills even in interactions with technology like ChatGPT.

While ChatGPT provides information based on text from the internet, it does not have the ability to critically evaluate and understand the credibility of theories. However, it is possible that students will be able to build their knowledge independently simply by using their interactions with ChatGPT. It should be noted that there is no guarantee that the knowledge they build will be justified true belief. Therefore, the presence of ChatGPT cannot provide justification of truth in the didactic tetrahedron concept. Because the processes of diffusion and acquisition cannot be fulfilled. Its role in didactics remains and will always be as a tool that, if substituted into the didactic triangle, is likely to be on the coordinates within the pedagogical didactic anticipation (ADP) line.

Several factors reinforce that ChatGPT cannot stand independently in the didactic concept: (1) ChatGPT relies on testimonial information taken from the internet, which includes various sources, both reliable and unreliable (Lund and Wang, 2023). It does not have the ability to distinguish between the two. This limitation poses the risk of spreading false information or unverified claims, as ChatGPT lacks the characteristics of critical evaluation skills; (2) ChatGPT, on the other hand, does not have the ability to ask targeted and in-depth questions that stimulate critical thinking and encourage individuals to reflect on their own perspectives (Bishop, 2023; Yu, 2023); (3) ChatGPT does not have the capacity to engage in dialectical exchanges and refine concepts collaboratively with students (Loos et al., 2023); (4) ChatGPT, as an AI language model, does not have the capacity to guide students in the same way, limiting the depth of critical self-reflection (Loos et al., 2023); (5) ChatGPT cannot engage in inductive reasoning processes, which are essential for developing high-level thinking skills (Echenique, 2023); (6) ChatGPT cannot engage in dialectical exchanges or provide the same level of challenge and intellectual development (Echenique, 2023).

4 Conclusion

In comparing the mathematical performance among three groups of students, significant differences in mathematical performance were observed. The group that solely relied on ChatGPT achieved the lowest average scores, with statistical results indicating a difference from the other two groups. In contrast, the group receiving instruction with the assistance of ChatGPT and the group receiving instruction solely from the instructor both achieved higher average scores, with no statistical difference in their mathematical performance. Therefore, this research underscores that the use of ChatGPT as a teaching tool in mathematics instruction has the potential to enhance student performance, but the role of the instructor remains crucial in delivering in-depth instruction. While ChatGPT can provide information, explanations, and support, the combination of human expertise and ChatGPT's capabilities holds greater potential for improving students' understanding of mathematics.

As a form of "extended cognition," the utilization of ChatGPT in mathematics education urgently needs development to ensure its optimal and appropriate use. The popularity of ChatGPT has reached significant levels, prompting researchers and developers to prioritize efforts in ensuring its effectiveness in the learning process rather than focusing solely on marketing aspects to attract students' interest. The

findings from this research show that students have a positive perception of ChatGPT's use in learning and enhancing their understanding of mathematical concepts. They view ChatGPT as a valuable tool for assisting them in composing papers or essays related to mathematics and understanding complex concepts. However, concerns exist regarding excessive dependence on ChatGPT, which could diminish students' ability to solve problems independently. There is a dual perspective among students, with some considering ChatGPT a valuable asset, while others worry about ethical implications and potential hindrances to intellectual development. Despite these concerns, students recognize the value of ChatGPT as a learning aid but emphasize the importance of its controlled use by instructors and mature ethical considerations.

The incorporation of ChatGPT does not provide valid justification for the didactical tetrahedron in the context of didactics. Although ChatGPT has the potential to provide information and answer questions, several reasons indicate that ChatGPT cannot independently stand as a valid component in didactic contexts. First, ChatGPT relies on information from the internet without the ability to distinguish between reliable and unreliable sources, potentially spreading inaccurate or unverified information. This limitation does not align with the principle of strong justification in building knowledge. Second, ChatGPT cannot pose deep questions that stimulate critical thinking and cannot engage in dialectical exchanges that promote reflection and intellectual development. Third, ChatGPT lacks the capacity to guide and provide human-like guidance to students, which is necessary for developing deep understanding and high-level thinking skills. Fourth, ChatGPT cannot engage in inductive reasoning, which is crucial for developing critical thinking skills. Therefore, while the use of ChatGPT as a teaching aid in mathematics education, supplemented by instructor guidance, is worthy of further investigation, its use in the learning process cannot validly replace the role of instructors or deep and reflective human interactions in constructing meaningful knowledge. Consequently, the inclusion of ChatGPT cannot provide valid justification for the didactical tetrahedron in the didactic context.

The findings of this research carry significant implications for teaching practices in university settings. Instructors continue to play a pivotal role in delivering in-depth and personalized mathematics instruction. While ChatGPT can offer supplementary support, human-led instruction remains irreplaceable in providing contextual explanations, discussing complex concepts, and stimulating questions and discussions. Overall, the use of ChatGPT as a teaching aid in mathematics instruction demonstrates significant potential for enhancing student performance. However, this research also emphasizes that the use of ChatGPT needs to be integrated with deep instructor-led instruction to be effective. Future research can explore the development of more advanced language models, optimization strategies for using ChatGPT in education, and the specific roles of instructors in leveraging this teaching aid effectively.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical approval was not required for the studies involving humans because this type of research, in Indonesia, there are no ethical committee available yet. The ethical committee is commonly used in medical research. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

DD: Writing – original draft, Formal analysis. AH: Writing – original draft, Conceptualization, Data curation. SS: Resources, Software, Writing – original draft. DS: Formal analysis, Supervision, Validation, Writing – review & editing. LM: Project administration, Software, Writing – original draft. TC: Writing – review & editing, Supervision. LF: Project administration, Resources, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Several authors have received educational scholarships from the Indonesian

government through the Education Fund Management Institution (LPDP). This research is partially funded by the LPDP.

Acknowledgments

The authors would like to express their gratitude to the Education Fund Management Institution (Lembaga Pengelola Dana Pendidikan—LPDP) under Ministry of Finance of the Republic of Indonesia as sponsor for their Masters' and Doctoral studies. Authors also would like to express their sincere gratitude to all individuals and institutions whose support and contributions made this research possible.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Adiguzel, T., Kaya, M. H., and Cansu, F. K. (2023). Revolutionizing education with AI: exploring the transformative potential of ChatGPT. *Contemp. Educ. Technol.* 15, 1–13. doi: 10.30935/cedtech/13152
- Alam, A. (2022). Employing adaptive learning and intelligent tutoring robots for virtual classrooms and smart campuses: reforming education in the age of artificial intelligence. In *Advanced computing and intelligent technologies: Proceedings of ICACIT 2022* (pp. 395–406). Springer.
- Aldon, G., Cusi, A., Schacht, F., and Swidan, O. (2021). Teaching mathematics in a context of lockdown: A study focused on teachers' praxeologies. *Educ. Sci.* 11, 1–21. doi: 10.3390/educsci11020038
- Alneyadi, S., Abulibdeh, E., and Wardat, Y. (2023). The Impact of Digital Environment vs. Traditional Method on Literacy Skills: Reading and Writing of Emirati Fourth Graders. *Sustainability*, 15:3418.
- Apostolopoulos, I. D., Tzani, M., and Aznaouridis, S. I. (2023). ChatGPT: ascertaining the self-evident. The use of AI in generating human knowledge. *Comput. Therm. Sci.* 1, 1–20. doi: 10.48550/arXiv.2308.06373
- Aydın, Ö., and Karaarslan, E. (2023). Is ChatGPT leading generative AI? What is beyond expectations? *Acad. Platform J. Engineer. Smart Syst.* 11, 118–134. doi: 10.21541/apjess.1293702
- Baidoo-Anu, D., and Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): understanding the potential benefits of chatgpt in promoting teaching and learning. *J. AI* 7, 52–62. doi: 10.61969/jai.1337500
- Bishop, L. (2023). *Can ChatGPT Think Like a Lawyer?* A Socratic Dialogue.
- Biswas, S. S. (2023). Role of Chat GPT in Public Health. *Ann Biomed Eng.* 51, 868–869. doi: 10.1007/s10439-023-03172-7
- Bosch, M. (2015). Doing research within the anthropological theory of the didactic: the case of school algebra. In *Selected regular lectures from the 12th international congress on mathematical education* (pp. 51–69).
- Brinkley-Etz Korn, K. E. (2018). Learning to teach online: measuring the influence of faculty development training on teaching effectiveness through a TPACK lens. *Internet High. Educ.* 38, 28–35. doi: 10.1016/j.iheduc.2018.04.004
- Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht: Kluwer Academic Publishers.
- Cao, Y., Zhang, S., Chan, M. C. E., and Kang, Y. (2021). Post-pandemic reflections: lessons from Chinese mathematics teachers about online mathematics instruction. *Asia Pac. Educ. Rev.* 22, 157–168. doi: 10.1007/s12564-021-09694-w
- Coskun, H. (2023). The power of ChatGPT: the breakthrough role of the language model in engineering education. *Innovative Res. Engineer.* 71–100. doi: 10.59287/ire.507
- Creswell, J. W. (2012). Educational research: planning, conducting, and evaluating quantitative and qualitative research. *Educ. Res.* 4:237.
- Deiana, G., Dettori, M., Arghittu, A., Azara, A., Gabutti, G., and Castiglia, P. (2023). Artificial intelligence and public health: evaluating chat GPT responses to vaccination myths and misconceptions. *Vaccine* 11, 1–13. doi: 10.3390/vaccines11071217
- Dieronitou, I. (2014). The ontological and epistemological foundations of qualitative and quantitative approaches to research. *Int. J. Econ., Commerce and Manag. United* 11, 1–17.
- Echenique, C. R. (2023). Education in Latin America, what could we do? *Revista Boletín Redipe* 12, 37–60. doi: 10.36260/rbr.v12i4.1951
- Elbanna, S., and Armstrong, L. (2023). *Exploring the integration of ChatGPT in education: adapting for the future*. Management & Sustainability: an Arab Review.
- Eriksson, H., and Larsson, N. (2023). Chatting up the grade an exploration on the impact of ChatGPT on self-study experience in higher education (Umeå University). Umeå University. Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-211263%0A>
- Fischer, E., and Guzel, G. T. (2023). The case for qualitative research. *J. Consum. Psychol.* 33, 259–272. doi: 10.1002/jcpy.1300
- Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht: Kluwer Academic Publishers.
- Gautam, A., Chirputkar, A., and Pathak, P. (2022). Opportunities and challenges in the application of artificial intelligence-based technologies in the healthcare industry.

International interdisciplinary humanitarian conference for sustainability, IIHC 2022-proceedings, 18(1), 1521–1524. MDPI.

Geis, J. R., Brady, A., Wu, C. C., Spencer, J., Ranschaert, E., Jaremko, J. L., et al. (2019). Ethics of artificial intelligence in radiology: summary of the joint European and north American multisociety statement. *Insights Imag* 10, 101–440. doi: 10.1186/s13244-019-0785-8

Ghavifekr, S., and Rosdy, W. A. W. (2015). Teaching and learning with technology: effectiveness of ICT integration in schools. *Int. J. Res. Educ. Sci.* 1, 175–191. doi: 10.21890/ijres.23596

Gong, B., Nugent, J. P., Guest, W., Parker, W., Chang, P. J., Khosa, F., et al. (2019). Influence of artificial intelligence on Canadian medical students' preference for radiology specialty: A national survey study. *Acad. Radiol.* 26, 566–577. doi: 10.1016/j.acra.2018.10.007

Gupta, A., and Elby, A. (2011). Beyond epistemological deficits: dynamic explanations of engineering students' difficulties with mathematical sense-making. *Int. J. Sci. Educ.* 33, 2463–2488. doi: 10.1080/09500693.2010.551551

Haleem, A., Javaid, M., and Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *Bench Council Transactions on Benchmarks, Standards and Evaluations* 2:100089. doi: 10.1016/j.tbench.2023.100089

Hassan, A. M., Nelson, J. A., Coert, J. H., Mehrara, B. J., and Selber, J. C. (2023). Exploring the Potential of Artificial Intelligence in Surgery: Insights from a Conversation with ChatGPT. *Annals of surgical oncology* 30, 3875–3878.

Hassani, H., and Silva, E. S. (2023). The role of chatGPT in data science: how AI-assisted conversational interfaces are revolutionizing the field. *Big Data and Cognitive Computing* 7:62. doi: 10.3390/bdcc7020062

Heggart, K. R., and Yoo, J. (2018). Getting the most from google classroom: A pedagogical framework for tertiary educators. *Australian J. Teacher Educ.* 43, 140–153. doi: 10.14221/ajte.2018v43n3.9

Javaid, M., Haleem, A., and Singh, R. P. (2023). ChatGPT for healthcare services: an emerging stage for an innovative perspective. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* 3:100105. doi: 10.1016/j.tbench.2023.100105

Jeon, J., and Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Educ. Inf. Technol.* 28, 15873–15892. doi: 10.1007/s10639-023-11834-1

Jukić Matić, L., and Glasnović Gracin, D. (2016). Das Schulbuch als Artefakt in der Klasse: Eine Studie vor dem Hintergrund des sozioididaktischen Tetraeders. *J. Für Mathematik-Didaktik* 37, 349–374. doi: 10.1007/s13138-016-0091-7

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* 103:102274. doi: 10.1016/j.lindif.2023.102274

Kirk, R. E. (2008). *Statistics: An introduction (fifth)*. United States of America: Michele Sordi.

Kirk, R. E. (2009). Experimental design. *Sage handbook of quantitative methods in psychology*, 23–45.

Langer, J. A. (2011). *Envisioning knowledge: Building literacy in the academic disciplines*. New York: Teachers College Press.

Lecler, A., Duron, L., and Soyer, P. (2023). Revolutionizing radiology with GPT-based models: current applications, future possibilities and limitations of ChatGPT. *Diagn. Interv. Imaging* 104, 269–274. doi: 10.1016/j.diii.2023.02.003

Lee, C. J. G. (2012). Reconsidering constructivism in qualitative research. *Educ. Philos. Theory* 44, 403–412. doi: 10.1111/j.1469-5812.2010.00720.x

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Educ. Sci.* 13:410. doi: 10.3390/educsci13040410

Loos, E., Gröpler, J., and Goudeau, M. L. S. (2023). Using ChatGPT in education: human reflection on ChatGPT's self-reflection. *For. Soc.* 13, 1–18. doi: 10.3390/soc13080196

Lund, B. D., and Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library Hi Tech News* 40, 26–29. doi: 10.1108/LHTN-01-2023-0009

Marshall, P. A., Adebamowo, C. A., Adeyemo, A. A., Ogundiran, T. O., Vekich, M., Strenski, T., et al. (2006). Voluntary participation and informed consent international genetic research. *Am. J. Public Health* 96, 1989–1995. doi: 10.2105/AJPH.2005.076232

Mattas, P. S. (2023). ChatGPT: A study of AI language processing and its implications. *Int. J. Res. Public. Rev.* 4, 435–440. doi: 10.55248/gengpi.2023.4218

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of chatGPT towards lifelong learning. *SSRN Electron. J.* doi: 10.2139/ssrn.4354422

Mitrović, S., Andreoletti, D., and Ayoub, O. (2023). ChatGPT or human? Detect and explain. Explaining decisions of machine learning model for detecting short ChatGPT-generated text. *Comput. Therm. Sci.* 1, 1–11. doi: 10.48550/arXiv.2301.13852

Morrison, D. R. (2010). *The Cambridge companion to socrates*. Cambridge: Cambridge University Press.

Novita, R., and Herman, T. (2021). Using technology in young children mathematical learning: A didactic perspective. *J. Phys. Conf. Ser.* 1957, 012013–012010. doi: 10.1088/1742-6596/1957/1/012013

Olive, J., Makar, K., Hoyos, V., Kor, L. K., Kosheleva, O., and Sträßer, R. (2010). Mathematical knowledge and practices resulting from access to digital technologies. *New ICMI study series* 13, 133–177. doi: 10.1007/978-1-4419-0146-0_8

Oviedo-Trespalcacios, O., Peden, A. E., Cole-Hunter, T., Costantini, A., Haghani, M., Rod, J. E., et al. (2023). The risks of using Chatgpt to obtain common safety-related information and advice. *SSRN Electron. J.* 167:106244. doi: 10.2139/ssrn.4370050

Owens, T. J. (1950). Socratic method and critical philosophy: selected essays. *Thought* 25, 153–154. doi: 10.5840/thought195025134

Pavlik, J. V. (2023). Collaborating with ChatGPT: considering the implications of generative artificial intelligence for journalism and media education. *Journalism and Mass Communication Educator* 78, 84–93. doi: 10.1177/10776958221149577

Pedró, F. (2019). Artificial intelligence in education: challenges and opportunities for sustainable development. Unesco, 46. Retrieved from <https://en.unesco.org/themes/education-policy>

Phuong, T. T., Cole, S. C., and Zarestky, J. (2018). A systematic literature review of faculty development for teacher educators. *High. Educ. Res. Dev.* 37, 373–389. doi: 10.1080/07294360.2017.1351423

Prediger, S., Roesken-Winter, B., and Leuders, T. (2019). Which research can support PD facilitators? Strategies for content-related PD research in the three-tetrahedron model. *J. Math. Teach. Educ.* 22, 407–425. doi: 10.1007/s10857-019-09434-3

Rahman, M. S. (2016). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language “testing and assessment” research: A literature review. *J. Educ. Learn.* 6:102. doi: 10.5539/jel.v6n1p102

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. doi: 10.1016/j.iotcps.2023.04.003

Rees, T. (2022). Non-human words: on GPT-3 as a philosophical laboratory. *Daedalus* 151, 168–182. doi: 10.1162/DAED_a_01908

Rehana, H., Çam, N. B., Basmacı, M., He, Y., Özgür, A., and Hur, J. (2023). Evaluation of GPT and BERT-based models on identifying protein-protein interactions in biomedical text. *ArXiv Preprint ArXiv* 10, 1–25. doi: 10.48550/arXiv.2303.17728

Remillard, J. T., and Heck, D. J. (2014). Conceptualizing the curriculum enactment process in mathematics education. *ZDM-Math. Educ.* 46, 705–718. doi: 10.1007/s11858-014-0600-4

Rezat, S., and Sträßer, R. (2012). From the didactical triangle to the socio-didactical tetrahedron: artifacts as fundamental constituents of the didactical situation. *ZDM-Math. Educ.* 44, 641–651. doi: 10.1007/s11858-012-0448-4

Ruthven, K. (2009). Towards a naturalistic conceptualisation of technology integration in classroom practice: the example of school mathematics. *Éducation et Didactique* 9, 131–159. doi: 10.4000/educationdidactique.434

Ruthven, K. (2012). The didactical tetrahedron as a heuristic for analysing the incorporation of digital technologies into classroom practice in support of investigative approaches to teaching mathematics. *ZDM-Math. Educ.* 44, 627–640. doi: 10.1007/s11858-011-0376-8

Safdar, N. M., Banja, J. D., and Meltzer, C. C. (2020). Ethical considerations in artificial intelligence. *Eur. J. Radiol.* 122:108768. doi: 10.1016/j.ejrad.2019.108768

Shaikh, S., Yayilgan, S. Y., Klimova, B., and Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European J. Investigation in Health, Psychol. Educ.* 13, 1937–1960. doi: 10.3390/ejihpe13090140

Singh, J., Steele, K., and Singh, L. (2021). Combining the best of online and face-to-face learning: hybrid and blended learning approach for covid-19, post vaccine, & post-pandemic world. *J. Educ. Technol. Syst.* 50, 140–171. doi: 10.1177/00472395211047865

Sit, C., Srinivasan, R., Amlani, A., Muthuswamy, K., Azam, A., Monzon, L., et al. (2020). Attitudes and perceptions of UK medical students towards artificial intelligence and radiology: a multicentre survey. *Insights Imag* 11, 14–16. doi: 10.1186/s13244-019-0830-7

Steinbring, H. (2005). Analyzing mathematical teaching-learning situations - the interplay of communicational and epistemological constraints. *Educ. Stud. Math.* 59, 313–324. doi: 10.1007/s10649-005-4819-4

Straesser, R. (2007). Didactics of mathematics: more than mathematics and school! *ZDM-Int. J. Mathematics Educ.* 39, 165–171. doi: 10.1007/s11858-006-0016-x

Sun, G. H., and Hoelscher, S. H. (2023). The ChatGPT storm and what faculty can do. *Nurse Educ.* 48, 119–124. doi: 10.1097/NNE.0000000000001390

Suryadi, D. (2019a). *Landasan Filosofis Penelitian DDR*. Bandung: Gapura Press.

Suryadi, D. (2019b). *Penelitian Desain Didaktis (DDR) dan Implementasinya*. Bandung: Gapura Press.

- Tall, D. (1986). Using the computer as an environment for building and testing mathematical concepts: a tribute to richard skemp. *Honour of Richard Skemp*, 21–36.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., et al. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learn. Environ.* 10, 1–10. doi: 10.1186/s40561-023-00237-x
- To, J., and Carless, D. (2016). Making productive use of exemplars: peer discussion and teacher guidance for positive transfer of strategies. *J. Furth. High. Educ.* 40, 746–764. doi: 10.1080/0309877X.2015.1014317
- Troussas, C., Krouska, A., and Virvou, M. (2020). Using a multi module model for learning analytics to predict learners' cognitive states and provide tailored learning pathways and assessment. *Intelligent systems reference library* 158, 9–22. doi: 10.1007/978-3-030-13743-4_2
- Tuli, F. (2011). The basis of distinction between qualitative and quantitative research in social science: reflection on ontological, epistemological and methodological perspectives. *Ethiopian J. Educ. Sci.* 6, 97–108. doi: 10.4314/ejesc.v6i1.65384
- Vincent-Lancrin, S., and van der Vlies, R. (2020). *Trustworthy artificial intelligence (AI) in education: promises and challenges*. OECD education working papers, No. 218. In OECD Publishing. France.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Wardat, Y., Tashtoush, M. A., AlAli, R., and Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia J. Math. Sci. Technol. Educ.* 19, em2286–em2218. doi: 10.29333/ejmste/13272
- Wittmann, J. (2023). Science fact vs science fiction: A ChatGPT immunological review experiment gone awry. *Immunol. Lett.* 256–257, 42–47. doi: 10.1016/j.imlet.2023.04.002
- Yang, Y. T. C., and Wu, W. C. I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking; learning motivation: A year-long experimental study. *Comp. Educ.* 59, 339–352. doi: 10.1016/j.compedu.2011.12.012
- Yu, H. (2023). Reflection on whether chat GPT should be banned by academia from the perspective of education and teaching. *Front. Psychol.* 14, 1–12. doi: 10.3389/fpsyg.2023.1181712
- Zhang, X., Shah, J., and Han, M. (2023). ChatGPT for fast learning of positive Energy District (PED): A trial testing and comparison with expert discussion results. *Buildings* 13:1392. doi: 10.3390/buildings13061392



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Freddy Marín-González,
Costa University Corporation, Colombia
Kerry Shephard,
University of Otago, New Zealand

*CORRESPONDENCE

Claudia Camacho-Zuñiga
✉ claudia.camacho@tec.mx

RECEIVED 27 June 2023

ACCEPTED 08 January 2024

PUBLISHED 25 January 2024

CITATION

Cuevas-Cancino MdIO, Peña-Becerril M,
Mondragon-Estrada E and
Camacho-Zuñiga C (2024) Incorporating
vertical collaboration to address sustainable
development goals: the Monarch Route
Project.

Front. Educ. 9:1246889.

doi: 10.3389/feduc.2024.1246889

COPYRIGHT

© 2024 Cuevas-Cancino, Peña-Becerril,
Mondragon-Estrada and Camacho-Zuñiga.
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.

Incorporating vertical collaboration to address sustainable development goals: the Monarch Route Project

Maria de la O Cuevas-Cancino¹, Maritza Peña-Becerril¹,
Enrique Mondragon-Estrada¹ and Claudia Camacho-Zuñiga^{2*}

¹School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico, ²Institute for the Future of Education, School of Engineering and Sciences, Tecnológico de Monterrey, Monterrey, Mexico

Higher Education Institutions in their societal transformative role need to effectively integrate the UN's Sustainable Development Goals (SDGs) and the 2030 Agenda into their curricula. Our goal is to inspire and trigger the discussion in the academic community on learning environments for education for sustainable development. We report *The Monarch Route* (MRP), a project conducted at *Tecnológico de Monterrey* and awarded with the 2018 *Premios Latinoamerica Verde*. Engaging all the students enrolled in the different semesters of the Sustainable Development Engineering (SDE) program, it adopted a research-based learning methodology, which is highly flexible at setting the final deliverables and due dates. It also implemented horizontal (intra-class), vertical (inter-class), and interdisciplinary (multiple courses) student collaboration. In partnership with a non-governmental organization, the MRP aimed to reverse the extinction threat to the monarch butterfly while benefiting rural communities along the butterfly's migratory route. With an exploratory action research approach, we analyzed the data gathered 2 months after the end of the courses: A survey on the perceived learning outcomes (87% confidence level) and a letter for future classes sharing their experience on the MRP, both optional. The analysis included a mixed methodology, i.e., quantitative and qualitative, including text mining, which evidenced that the participants strengthened both, transverse competences (collaboration, citizenship, communication, and critical thinking) and disciplinary competences (subject matter). The MRP is innovative for its dynamic structure that resembles the interconnectedness of real professional practice, scientific research, and environmental conservation efforts. In addition to the disciplinary application and societal relevance, the instructional design of MRP reflects that contributing to 2030 Agenda requires a multi-faceted approach, involving educational institutions, community partnerships, and interdisciplinary collaboration. Traditional education no longer fulfills our needs to face environmental challenges; academics, institutions, and governments are encouraged to implement policies and strategies to educate for a sustainable and better future for humanity and the planet.

KEYWORDS

vertically integrated project, research-based learning, sustainable development goals in higher education, sustainability competence, sustainability learning, education for 2030 Agenda, higher education

1 Introduction

The urgency needed to accelerate progress toward the Sustainable Development Goals (SDGs) (UNESCO, 2015) is undeniable. Global Sustainable Development Report 2023 states that the required transformations are possible and inevitable and invites us to embrace them. Since the Global Sustainable Development Report was published in 2019, challenges have multiplied and intensified, causing the world to fall short of achieving the SDGs. Despite the progress made in areas such as the deployment of zero carbon technologies, several factors, such as the ongoing Ukraine–Russia conflict, inflation, cost-of-living crisis, and environmental, economic, and societal distress, have disrupted overall progress toward the 2030 Agenda and the SDGs (United Nations, 2023).

Our future relies on multiple securities, including geopolitical, energy, climate, water, food, and social security. Thus, strategies to embrace transformations should uphold the principles of solidarity, equity, and well-being, harmoniously with nature. As the world rapidly evolves, there is an increased urgency to build momentum, solidify solidarity, and speed up progress on the SDGs (United Nations, 2023).

The scope of Sustainable Development expands beyond the environmental aspect. It includes economic, social, and political dimensions well identified by the wide variety of topics included in the SDGs. In this regard, achieving the 2030 Agenda demands the collaboration of different societal entities, including governments, non-governmental organizations (NGOs), civil society, and educational institutions (United Nations, 2023). Education is, in fact, a fundamental component represented in SDG 4, quality education (UNESCO, 2015). Then, far from a mere environmental education, educational institutions must integrate sustainability into all levels and forms of education, influencing curriculum design, teaching methodologies, learning environment, and the entire educational experience (UNESCO and UNESCO Global Independent Expert Group on the Universities and the 2030 Agenda, 2022).

Higher education institutions (HEIs) should lead society and invest time and resources as effectively as possible, taking a systematic and strategic approach to drive and accelerate transformations. HEIs must design learning environments to build sustainability concepts and practices from the whole curricula (UNESCO and UNESCO Global Independent Expert Group on the Universities and the 2030 Agenda, 2022). Including real-world contexts and societal impacts in those environments will encourage intellectual cooperation and cross-disciplinary collaborations. This cooperation can accelerate the generation and sharing of knowledge, cultivate understanding of the SDGs, and inspire innovation in strategies for sustainable development (UNESCO and UNESCO Global Independent Expert Group on the Universities and the 2030 Agenda, 2022). Additionally, by integrating SDG-related research and teaching into their curricula, HEIs can shape students with the necessary skills and knowledge to become citizens who contribute effectively to sustainable development in their future career and life (Mazur and Walczyna, 2022; Acevedo-Duque et al., 2023).

The *Premios Latinoamerica Verde* (Green Latin America Awards) is one of the most significant environmental awards in the Latin America region. Launched in 2013 in Ecuador, the awards aim to recognize, promote, and encourage best practices and innovative

projects related to environmental sustainability and social responsibility in Latin American and Caribbean countries. The awards offer a platform for the sharing of knowledge and ideas, promoting regional integration and cooperation toward a common goal: the development of a more sustainable and environmentally friendly Latin America. This is the way to encourage and motivate individuals, communities, and organizations to consider and act upon the environmental challenges facing the region (Premios Verdes Team, 2023).

The Monarch Route (MRP) was a university project awarded with the *Premios Latinoamerica Verde* in 2018 (Latinoamerica Verde, 2018). MRP followed a research-based learning (RBL) methodology and implemented a vertical collaboration with the entire student body enrolled in the Sustainable Development Engineering program of the Tecnológico de Monterrey, Campus Santa Fe, in 2017. It was conducted in conjunction with the *Fundación Nacional para la Conservación del Hábitat Boscoso de la Mariposa Monarca*, A.C. (FUNACOMM). It aims to stop and reverse the disappearance of the monarch butterfly by focusing on the development of agroecological gardens along its migratory route in Mexican territory. This was done to increase the availability of food and rest of the areas that directly aid in the conservation of the species. The project heavily relies on the use of technologies for monitoring and promoting environmental education within the communities located along the butterfly's migratory route.

Since the year 2000, various researchers and institutions interested in the subjects that have proposed different strategies for the conservation of the species, such as expanding the protection zone including conservation incentives (Anderson and Brower, 1996; Andrade and Rhodes, 2012). However, by 2017, the population of butterflies arriving in Mexico had decreased by 95% (Agrawal, 2017). The main threats have been identified as follows: deforestation due to illegal logging in various areas of the migratory route and the reserves located in Mexico; reduction in breeding areas due to land use changes in some places; decrease in the primary food source, known as *Asclepias* or *algodoncillo*, for the butterfly larvae due to the use of herbicides; and finally, climate change (Brower et al., 2002, 2012).

From its instructional design, the MRP aimed to impact several SDGs and provide students with firsthand experience in activities focused on achieving 2030 Agenda and the SDGs in Mexico. The MRP sought to give undergraduate students the opportunity to understand, design, and implement actions in support of the SDGs. It aimed to guide them on their path to becoming global citizens, enhancing their understanding of global and local problems, and inspiring them to continue implementing actions for a more sustainable future.

The goal of this study is to trigger the discussion among the academic community on instructional design and learning environments that actively contribute to achieving the 2030 Agenda. We present an instructional practice implemented in the Monarch Route Project MRP, an award-winning university initiative that engaged students from the sustainable engineering program in a research-based learning experience. We describe an educational setting involving interdisciplinary cooperation, incorporating both horizontal (peer-to-peer, same classes) and vertical (students from different classes) collaborative work, and addressing environmental

and productive challenges for a Mexican community. By analyzing the framework of the MRP, activities, and the perceptions of the participants, the current study highlights the effective strategies and pedagogical practices that contributed to the success of the project in fostering real-world impact and increasing the students' competences for a sustainable future.

2 Pedagogical framework

2.1 Education for sustainability

Ecological disasters highlight that the education and traditional programs we currently offer are not up to the social and environmental challenges we face in this era (Frisk and Larson, 2011; Fernández, 2023). There is an urgent need to adopt measures to address the increasing environmental risks of this era; our main problem is not the lack of solutions but the perspective from which current crises are gestated and understood (Fernández, 2023). We assume that knowledge automatically leads to action; however, we must go beyond ecological knowledge when educating for true transformation (Frisk and Larson, 2011; UNESCO and UNESCO Global Independent Expert Group on the Universities and the 2030 Agenda, 2022).

Education for and of the future must avoid depriving millions of species of their right to flourish, impoverishing future generations (Kopnina, 2020); such education for sustainability must be imparted now. Throughout history, education has played a crucial role in solving social problems, but in the Anthropocene, it is not only responsible for disseminating information but also for developing sustainability competence in students, reshaping their cognitive paradigms (Fernández, 2023). Ecopedagogy, ecocentric and indigenous learning, is a response to the educational needs of the Anthropocene (Kopnina, 2020).

In the Anthropocene, education must develop effective approaches for transformative actions, which implies that alternative ways of competence strengthening are essential for effectively educating for sustainability (Frisk and Larson, 2011; UNESCO and UNESCO Global Independent Expert Group on the Universities and the 2030 Agenda, 2022). A critical and enhanced education for sustainability involves developing deep competences, as Wamsler (2020) asserts that it is an internal transformation. This explains the interconnection between education for sustainability and human development (Agbedahin, 2019) and the urgent need for interdisciplinary, systemic, and comprehensive pedagogical approaches (Wamsler, 2020). We cannot face the Anthropocene with the same educational, social, and economic tools we applied in the preceding era.

2.2 Research-based learning

The current study assumes that the transformative actions required in Education for Sustainability are attainable through research-based learning (RBL). It is a teaching and learning strategy that takes the student from traditional passive learning to a leading role in what is to come (Boyer Commission on Educating Undergraduates in the Research University, 1998). The student goes from being a receiver of information to becoming a promoter and

producer of knowledge. This change takes place gradually over time because RBL allows the student to learn by means of discovery until they can generate and transfer knowledge by themselves.

Among the competences developed by RBL, we can identify open-mindedness, creative and critical thinking, research strategies, information analysis, and information source valuation. To obtain these, the instructor must design activities that allow the student to identify, analyze, and solve problems (Boyer Commission on Educating Undergraduates in the Research University, 1998). In addition, the relationship between research and teaching generates immediate learning benefits in students due to their mutual complementarity (Nehme, 2012; López-Gómez, 2015).

According to the Self-Determination Theory (Ryan and Deci, 2000), there are four RBL strategies that foster student motivation as follows:

- **Choice:** Empowering students with the opportunity to make decisions enhances their ownership of the learning process. Choice has been linked to increased effort, improved performance quality, and a higher likelihood of project completion, making it a key component of successful learning.
- **Challenge:** Engaging students with demanding tasks, even those that push beyond their perceived abilities, can heighten their involvement. However, this strategy requires careful management. A challenge with an adequate level of complexity encourages students and overcoming it can enhance their sense of competence and satisfaction. Conversely, an excessively difficult or unattainable challenge may discourage student interest.
- **Collaboration:** Humans are inherently social, and collaborative work can stimulate learning. Through interaction, students can learn by discussing, clarifying ideas, and receiving feedback, thereby collectively constructing knowledge. Collaboration is most effective when students undertake parallel or complementary activities that they perceive as relevant to the situation.
- **Control:** Encouraging students to manage their own learning can increase their engagement in the process. Self-directed learners tend to be more receptive and honest about their work, performance, and abilities. They are often more willing to acknowledge their mistakes and demonstrate greater resilience in the face of failure.

2.3 Vertical collaboration

For graduates, vertical collaboration is a competence as crucial as horizontal and interdisciplinary collaboration. Vertical learning is characterized as an educational experience that fosters a community of learners from diverse academic levels—ranging from bachelor's to master's degree students. This community-based learning model encourages peer feedback, learning from mentors and industry liaisons and sharing of thematic interests (Gómez Puente and Doulougeri, 2022).

A related concept, known as Vertically Integrated Projects (VIP), is a multidisciplinary, multi-semester, team-based learning structure that engages students in hands-on projects, which are applicable to

various contexts (Coyle et al., 2006; Sonnenberg-Klein and Coyle, 2022). VIP teams are composed of students from diverse academic disciplines and levels, from sophomores to PhD candidates, and these teams work on long-term projects, with undergraduate students having the opportunity to participate in a project for multiple semesters (Coyle et al., 2006).

VIP fosters the development of several key skills, including teamwork, interpersonal communication, public speaking, personal responsibility, time management, and problem-solving (Coyle et al., 2006; Morini et al., 2022). Coyle et al. (2006) underscore the significance of coherence, skill-based labor distribution, team-based problem-solving, successful collaboration, and the ability to work with individuals of diverse backgrounds and styles.

Notably, the VIP model emphasizes personal responsibility as a crucial complement to teamwork. Students in VIP settings have recognized the value of taking initiative, exceeding expectations, exhibiting self-discipline, and demonstrating responsibility. This finding is particularly significant as it addresses the limitations often associated with team-based instruction, as documented in the literature (Coyle et al., 2006).

3 Learning environment

The Monarch Route Project (MRP) was proposed as an innovative research-based learning strategy. Its uniqueness lies in its interdisciplinarity, horizontally and vertically integrated collaboration, social relevance, alignment with 2030 Agenda, and partnership with a non-governmental organization (NGO).

Initially, the MRP was chosen due to the relevance of its subject matter, making it a large enough project which allowed for the participation of different classes, courses, and semesters. Due to the scope of the MRP, it included several different areas of each participating class and course curricula, making it a relevant part of the students' graded academic work within their courses.

The MRP encouraged interdisciplinary collaboration, mirroring real-world work environments, where professionals from diverse fields and expertise collaborate to solve complex problems. The project involved 181 Sustainable Development Engineering students from Tecnológico de Monterrey, Campus Santa Fe. These students, from various classes and courses, collaborated closely during the August–December 2017 and January–May 2018 periods. Throughout the year-long project, the students worked in teams within their courses (horizontal) and, at the same time, in collaboration with other teams in different courses (interdisciplinary) and even different classes (vertical).

In an initial session, we introduced the project to all participants and assigned objectives based on their academic years (i.e., freshmen, sophomores, juniors, and seniors), informing the students of the weight of the project in their final grade.

The instructional design of the MRP focused on societal relevance and real-world impact. The undergraduates collaborated with FUNACOMM, an NGO dedicated to conserving the monarch butterfly, and several rural communities along the butterfly's migration route. This partnership brought real-world impact and applications to the MRP.

The project was built on three main objectives, aimed to address specific SDG:

- Developing a technological platform for easy monitoring of the species and the implementation of conservation tools (SDG 13: Climate action and SDG 15: Life on Land).
- Creating a teaching resource to raise environmental awareness among communities along the butterfly's route and contribute to broader conservation efforts (SDG 4: Quality Education, SDG 15: Life on Land and SDG 17: Partnerships for the Goals).
- Identifying the Mexican states most affected by conservation issues related to the monarch butterfly and establishing a network of agroecological gardens in these areas. (SDG 2: Zero Hunger, SDG 8: Decent Work and Economic Growth, SDG 12: Responsible Consumption and Production and SDG 15).

Despite being a requirement for the courses' final grade, the students were responsible for setting their own project deliverables and deadlines, with the purpose of fostering their project management skills and self-directed learning. The collaboration interactions and their relevance to each Sustainable Development Goal (SDG) are presented in Figure 1. This figure describes the collaboration among students of each class and course. The deadlines for partial and final project submissions were determined in agreement among teams from different classes and courses. Therefore, the details in the implementation and evaluation of this RBL instructional design will vary, depending on the issue to be addressed, the objectives of the project, the disciplines involved, the societal partnership (NGO), and the student's approach to achieve them.

4 Results

The instructional design of the Monarch Route Project (MRP) aimed to foster interdisciplinary, horizontal, and vertical collaboration under a research-based learning strategy for the understanding of Sustainable Development Goals (SDGs) and acting to enhance 2030 agenda. To analyze the students' perceived learning outcomes of this project, we conducted exploratory action research using a mixed methods approach.

Data were collected 2 months after the end of the courses using an optional survey and written final remarks within a "Letter to your classmates." The instrument, administered via Google Forms in Spanish and shown in Table 1, included five closed questions (0–4 Likert scale) and three open questions. Forty-five students responded to the survey, providing a confidence level of 87% with a 5% margin of error (García-García et al., 2013). The survey introduction informed students about the research purpose, the optional participation at any time along the instrument, the confidentiality of their responses, and the lack of impact of their answers in their evaluations and courses. In addition to requesting their informed consent to participate in the research, it is important to note that no personal data were collected in the survey.

According to gathered data, participation in the MRP led to students' perceived learning outcomes related to both transverse competences (collaboration, citizenship, communication, and critical thinking) and disciplinary competences (subject matter). Figure 2A illustrates their perceptions of the most significant learning outcomes from the MRP, with citizenship being the most frequently mentioned competence and project organization the least. In other words, students recognized the MRP as relevant from both a field of study

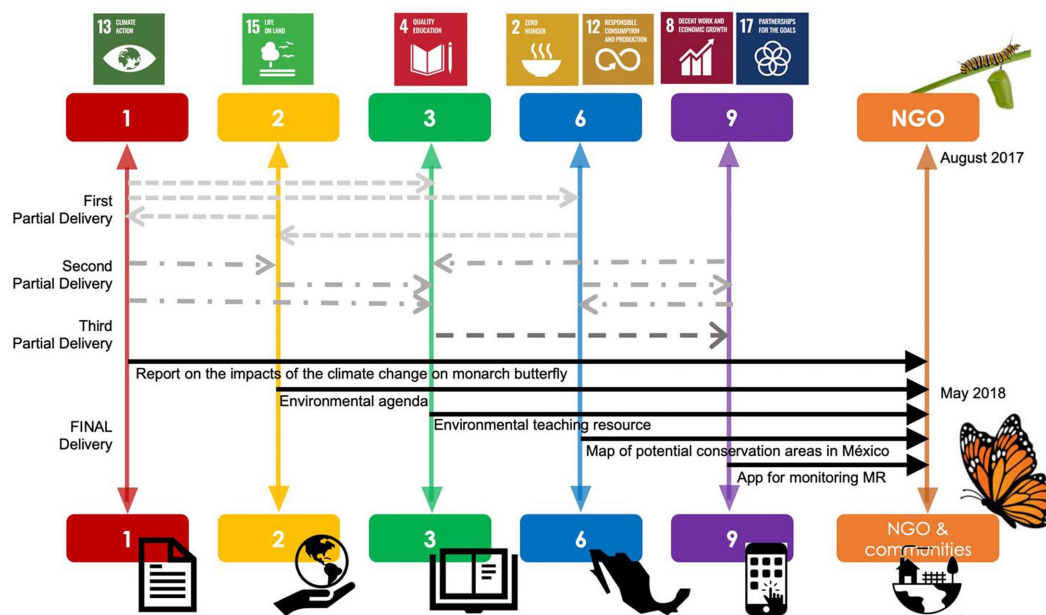


FIGURE 1

The interactions between each semester, specifying the different SDGs and delivery stages for each vertical, horizontal, and interdisciplinary collaboration in the Monarch Route Project, which was implemented with 181 undergraduates from Sustainable Development Engineering.

perspective (disciplinary competence such as sustainability) and a broader perspective (transverse competences such as human spirit, ethics, and citizenship). Figure 2B shows the students' perceptions of the extent to which their competences were strengthened through their work on the MRP. Notably, most competences were impacted "very much," with modes and medians of 4 for all competences, except for 'entrepreneurship' and 'leadership,' which had modes of 4 and medians of 3.

The perceived learning outcomes of the MRP were not only significant but also transferable to other contexts. Answers to questions Q3, Q4, Q6, and Q7 had modes and medians of 4 (very much). These results evidenced that students felt able to apply their learning outcomes in different contexts (Q3), highly recommended the MRP (Q4), found it valuable for the community (Q6), and felt empowered to act in favor of the environment (Q7).

The MRP was a motivating challenge for undergraduates from the characteristics of a Research-Based Learning (RBL) instructional design. Motivation is fundamental for significant self-managed learning outcomes; a motivated student maintains interest and commitment to the work, ensuring the achievement of the project's objectives. The MRP managed to maintain participants' motivation for an entire year, driven by the need for monarch butterfly conservation. Analyzing the MRP's characteristics from Ryan and Deci's Self-determination Theory (Ryan and Deci, 2000), we identified that motivation might be related to the following characteristics:

- The MRP gave students the opportunity to choose. The project's implementation instructions were broad, setting a clear and well-delineated objective without specifying the procedure or means. For instance, third-semester teams were asked to generate any

teaching resource related to environmental education that fostered the conservation of the monarch butterfly.

- The MRP presented students with a real, global challenge. This was due to the link with the social context surrounding the issue and its professional relevance.
- The MRP facilitated vertically and horizontally integrated collaboration, mirroring workplace environments. This meant students had to work with mates from their own class, other classes, and people in rural communities and FUNACOMM.
- The MRP allowed students to take control of their learning process. They could control the means and, in some cases, the final product to achieve the specified objective and the deadline submission.

Evaluating the project as a whole, over 95% of the students considered it useful for the community (Q6) and recommended it to others (Q4). The reasons for this recommendation are shown in Figure 2C, where the subject matter and the citizenship competence are the most mentioned.

The MRP's learning environment not only facilitated the participants, evidencing their sustainability competence, but also engaged them in continuing to act in favor of 2030 Agenda 2 months after the final submission. We categorized the answers of the students concerning the actions they were taking to solve the MRP problem (Q8). Our results demonstrate that after concluding the project, 82% of the participants effectively took actions as informed and active citizens by contributing to the monarch butterfly conservation. One of the top actions consisted of being responsible of their consumption (38% of the mentions), addressing SDG 12: Responsible Consumption and Production. Furthermore, 38% of the students tried to impact their close social network by sharing

TABLE 1 Survey used to gather data from the undergraduate participants in the Monarch Route project, 2 months after the final delivery to FUNACOM.

No.	Question	Possible answer	
Q1	Define in one sentence or phrase the most significant learning you obtained from this project.	Open-ended	
Q2	To what extent did the project contribute to the strengthening of your skills or competencies?		
	Problem-Solving	0-Not at all	4-Very much
	Self-directed Learning	0-Not at all	4-Very much
	Collaborative Work	0-Not at all	4-Very much
	Organization and Planning of a Project	0-Not at all	4-Very much
	Awareness of Your Environment	0-Not at all	4-Very much
	Leadership	0-Not at all	4-Very much
	Entrepreneurship	0-Not at all	4-Very much
	Empathy and Sensitivity Towards Environmental Issues	0-Not at all	4-Very much
	Reaffirming Your Career Vocation	0-Not at all	4-Very much
	Taking Action to Solve Problems in Your Environment	0-Not at all	4-Very much
	Oral and Written Communication	0-Not at all	4-Very much
	Global Perspective	0-Not at all	4-Very much
Q3	To what extent can you apply what you learned in this project to other projects?	0-Not at all	4-Very much
Q4	To what extent would you recommend this project?	0-Not at all	4-Very much
Q5	Why would you recommend this project?	Open-ended	
Q6	How useful do you believe this project was for the community?	0-Not at all	4-Very much
Q7	To what extent did this project empower you to participate in activities in favor of the environment?	0-Not at all	4-Very much
Q8	What are you doing today to contribute to the solution of the problem identified in the project?	Open-ended	

Five questions were in Likert scale (0—Not at all, 4—Very much) and the rest of the questions were open-ended.

their knowledge and promoting awareness, addressing SDG 4 education. Other SDGs involved were SDG 15: Life on Land (20%), SDG 13: Climate Action (9%), SDG 8: Decent Work and Economic Growth (7%), SDG 2: Zero Hunger (4%), and SDG 17: Partnerships for the Goals (2%).

After 2 months of the end of the course, i.e., after assigning final grades to the project's deliverables and the course, participants were invited to write an optional letter to be shared with students from incoming classes about the MRP. The letters, with no impact on their already registered grades, were collected through email; however, they were informed that disclosing their authorship was also optional. The specific instruction was: 'Write a letter to be shared with incoming classes telling your fellow students about the Monarch Route Project.' We obtained 22 letters and analyzed them using Text Mining, applying Voyant Tools (Sinclair and Rockwell, 2016), and ROBERTtuito (Pérez et al., 2022). The text analysis included the average sentence length and word count for each 'Letter to your classmates,' as well as vocabulary richness and sentiment. The final remarks had an average of 243 words, a vocabulary richness of 56%, and an average sentence length of 31 words.

The text analysis of our data reveals the MRP, designed under an RBL framework, develops the research skills of the students while also develops a sense of belonging and their positive attitude

to address related SDGs. Figure 3A presents a word cloud of the 25 most frequently mentioned words in the students' 'Letters to your classmates' about their experience with the MRP. Each term's size is proportional to its frequency. The terms 'project' (96), 'monarch' (70), 'butterfly' (55), 'route' (41), and 'ours' (40) were mentioned most frequently. The first four terms reflect the topic of the project: a project related to the monarch butterfly migration route. The term 'ours' signifies the students' sense of belonging and their commitment to addressing the threat of extinction faced by the monarch butterfly. This demonstrates that the students worked toward SDG 15: Life on Land and SDG 17: Partnerships for the Goals. Figure 3B is a diagram illustrating the 18 most frequently mentioned words and their interrelationships. The diagram's links reflect the concepts' articulation as mentioned by the participants. The nodes reveal subgroups related to disciplinary and transverse competences. For instance, the terms 'ours' and 'Mexico' relate to the MRP's region of analysis and the students' sense of belonging and awareness of the issue's impact. The term 'how,' which is associated with the students' ability to generate knowledge, strategies, and solutions independently, confirms the success of the RBL implementation.

The correlations between terms indicate that the students recognized improvements in their collaborative work, the RBL

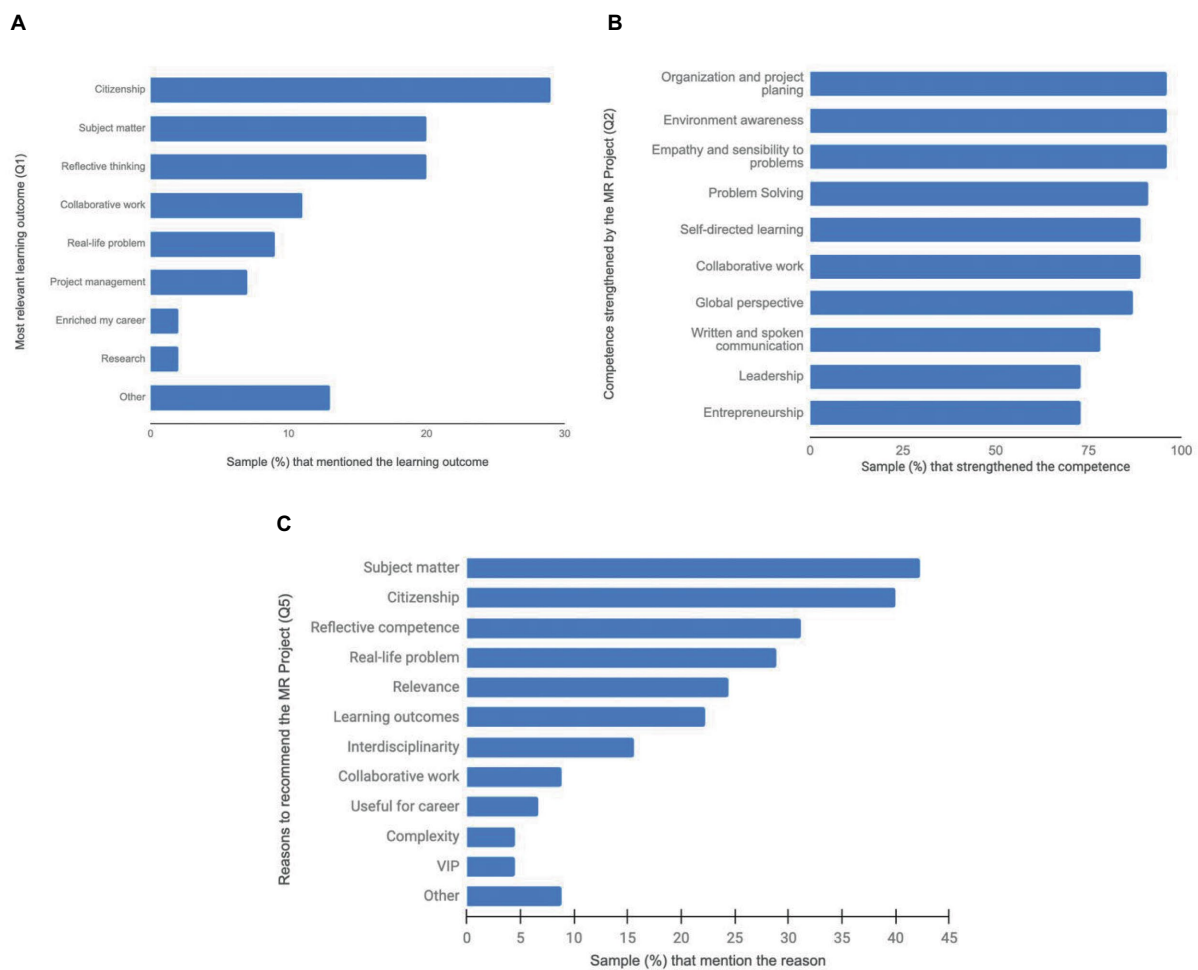


FIGURE 2

Participants in the Monarch Route Project (MRP) and their perception on (A) their most relevant learning outcome (Q1, open ended), on (B) the strengthening of specific competences (Q2, Likert scale), and on (C) the reasons to recommend this project (Q5, open ended). Answers to Q1 and Q5 were categorized by the authors while only answers 3 and 4—“very much” to Q2—are reported.

methodology, and their sustainability competence through the MRP. The vertically integrated collaboration is demonstrated by the term pairs ‘activity-collaborate’ and ‘collaborate-primary’ (Correlation Coefficient, $R = 0.9883$). The students perceived the relevance of the RBL methodology when they mentioned ‘research’ in sequence with ‘quality,’ ‘reliable,’ ‘source,’ and ‘fieldwork’ ($R > 0.9874$). The students’ significant learnings concerning the SDGs are evidenced by the word pairs mentioned in sequence, such as ‘establish-reserve’ ($R = 0.9688$, SDG 13 and SDG 15) and ‘citizens-ecosystem’ ($R = 0.9444$, SDG 13 and SDG 15), ‘establish-town’ ($R = 0.9688$, SDG 12 and SDG 17) and ‘collaborate-primary’ ($R = 0.9883$, SDG 17), ‘aware-economy’ ($R = 0.9456$, SDG 8 and SDG 12), and ‘knowledge-ecosystem’ and ‘disseminate-ecosystem’ ($R > 0.9444$, SDG 4).

Sentiment analysis of the respondents’ “Letter to your classmates” reflected to be mostly positive and neutral. We carried out sentiment analysis for “Letter to your classmates” with the pre-trained language model for user-generated text in Spanish RoBERTuito (Pérez et al., 2022), which is able to perform sentiment analysis, emotion analysis, irony detection, and hate speech detection. RoBERTuito categorized each answer to be positive, neutral, and negative, the prediction of

each category for each answer adds to one. Considering the category with the maximum probability, RoBERTuito labeled the “Letter to your classmates” to be 45% positive ($0.002 < \text{Probability} < 0.970$), 32% neutral ($0.017 < \text{Pr} < 0.661$), and 23% negative ($0.004 < \text{Pr} < 0.980$) (see examples in Table 2). The positive text shown focuses on the valuable learning experiences of the collaborative research work on a societal relevant project, and at the same time, the student’s motivation and engagement can be perceived. The neutral text reflects a description of the activity, highlighting the topic and the RBL strategy. Finally, the negative text can be associated with the students’ perception of the complexity and seriousness of the environmental problem associated with the extinction of the monarch butterfly.

5 Discussion

The Monarch Route Project (MRP) won the *Premios Latinoamérica Verde* 2018 award (Latinoamerica Verde, 2018). Its success is also supported by the students’ learning outcomes and transference perceptions, as well as on the high rate of recommendations they give. It was an innovative project from its

TABLE 2 Extracts from “Letter to your classmates,” a written text from the participants to the next cohort of students about the Monarch Route Project, the reflected sentiment, and Sustainable Development Goals they mentioned.

Sentiment (Probability)	Extract from the students “Letter to your classmates”	Student	Addressed SDG
Positive (Pr=0.970)	For me, one of the most important things to bring about change in the world is education, and that’s why I found this project so interesting. I was really motivated by the approach that the Monarch Route organization took with us and seeing how our work was going to be used in the real world. Also, I thought it was really cool how all of us from different years had a distinct role in this activity.	Student 11	SDG 4, SDG 17
Neutral (Pr=0.661)	Deforestation influences the death of the butterfly since they have nowhere to rest or eat. Because of this, for the Ecosystems and Biodiversity class, we did a project to monitor the route the monarchs follow when they migrate, and we researched their main causes of death to prevent them from dying and help them survive	Student 17	SDG 13, SDG 15
Negative (Pr=0.980)	The Monarch Butterfly project has the main objective of conserving the species of the monarch butterfly, its habitat, and its migration route. Currently, in Mexico there is not enough knowledge about the damage we are causing to the monarch butterfly route. Because of this, nowadays the route of this species is in danger, causing a big negative impact on the ecosystem.	Student 1	SDG 15, SDG 4

the entrepreneurship competence (social entrepreneurship and ecopreneurship) (Schaper, 2016).

- SDG 12: The MRP promoted responsible consumption and production practices by heightening awareness about the repercussions of human activities, particularly in agriculture, on the environment and fostering sustainable practices. Moreover, 2 months after the project’s deadline, most of the participants recognized taking effective action in this address.
- SDG 13: The MRP identified the effects of climate change on endangered species and laid the foundation for future action along the entire migration route of the monarch butterfly.
- SDG 15: The MRP additionally contributed to the preservation of terrestrial ecosystems, specifically those vital to the life cycle of the Monarch butterfly. Through education on the significance of these ecosystems and the promotion of conservation efforts, the MRP assists in safeguarding biodiversity and preventing ecosystem degradation and destruction.
- SDG 17: The MRP engaged in collaboration across diverse sectors and stakeholders, including Tecnológico de Monterrey, FUNACOMM, and rural communities. United, students, academia, and civil society will facilitate sustainable development and accomplish shared objectives.

The MRP encompassed species conservation understanding and recognition of its environmental and societal complexity and promoted self-directed learning; besides, students from Sustainable Development Engineering recognized the relevance of the project for their professional life and the strengthening of citizenship. Indeed, the MRP had a direct impact on the environmental issue, as FUNACOMM implemented all the proposals in rural communities along the monarch butterfly migration route. These features are internal learning motivators.

Although the students enrolled in any sustainable development course generally have prior interest in the subject matter, limiting the validity of our conclusions to our sample, it is noteworthy the MRP’s instructional design. The current global situation demands the collaboration of different societal agents, highlighting the relevance of fostering undergraduate participation in SDG-related projects, such as the MRP.

The MRP presented a challenge that enabled students to engage in interdisciplinary, horizontal, and vertical collaboration, choose, and govern their own learning process. These characteristics of the instructional design resemble the interconnectedness of real professional practice, scientific research, and environmental conservation efforts. It demonstrates that contributing to 2030 Agenda requires a multi-faceted approach, involving educational institutions, community partnerships, and interdisciplinary collaboration.

Following its implementation and subsequent recognition with the *Premios Latinoamérica Verde*, we prove the MRP as a valuable instructional model for Education for Sustainability. This action research illustrates how higher education can effectively integrate collaboration, real-world impact, and community engagement to empower students toward accomplishing 2030 Agenda. We invite the academic community to discuss on the value and achievements of such effort. The Monarch Route Project incorporated seven Sustainable Development Goals, adhering to a blueprint design for fostering peace and prosperity for both people and the planet, now and in the future.

Data availability statement

The datasets presented in this article are not readily available because the participants’ consent was provided under anonymity and confidentiality of their answers. Requests to access the datasets should be directed to claudia.camacho@tec.mx.

Ethics statement

Ethical approval was not required for the studies involving humans because at the moment of collecting the data, our Institution’s ethical requirement was to hold the signature of the informed consent from the participants. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was not obtained from the individual(s) for the publication of any

potentially identifiable images or data included in this article because published data are not individually identifiable.

Author contributions

MC-C contributed to instructional design and implementation of the Monarch Route Project. MP-B and CC-Z designed the survey and the “Letter to your classmates.” EM-E prepared the corpus and performed sentiment analysis. CC-Z wrote the first draft of the manuscript, elaborated the descriptive statistical analysis, and completed the text analysis (word cloud and graphs). MP-B wrote section 2.1 of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Acknowledgments

The authors would like to acknowledge the participation of *Fundación Nacional para la Conservación del Hábitat Boscoso de la*

Mariposa Monarca, A.C. (FUNACOMM) and the support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico in the production of this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Acevedo-Duque, Á., Jiménez-Bucarey, C., Prado-Sabido, T., Fernández-Mantilla, M. M., Merino-Flores, I., Izquierdo-Marín, S. S., et al. (2023). Education for sustainable development: challenges for postgraduate programmes. *Int. J. Environ. Res. Public Health* 20:1759. doi: 10.3390/ijerph20031759
- Agbedahin, A. V. (2019). Sustainable development, education for sustainable development, and the 2030 agenda for sustainable development: emergence, efficacy, eminence, and future. *Sustain. Dev.* 27, 669–680. doi: 10.1002/sd.1931
- Agrawal, A. (2017). Monarch population size over winter 2016–2017 announced. Insects on plants, chemical ecology, and coevolution. Available at: <http://www.eeb.cornell.edu/agrawal/2017/02/10/monarch-population-size-over-winter-2016-2017-announced/> (Accessed November 18, 2017).
- Anderson, J., and Brower, L. (1996). Freeze-protection of overwintering monarch butterflies in Mexico: critical role of the forest as a blanket and an umbrella. *Ecol. Entomol.* 21, 107–116. doi: 10.1111/j.1365-2311.1996.tb01177.x
- Andrade, G. S. M., and Rhodes, J. (2012). Protected areas and local communities: an inevitable partnership toward successful conservation strategies? *Ecol. Soc.* 17. Available at: <http://www.jstor.org/stable/26269207>
- Boyer Commission on Educating Undergraduates in the Research University (1998). Reinventing undergraduate education: A blueprint for America's research universities State University of New York at Stony Brook NY for the Carnegie Foundation for the Advancement of Teaching. Stony Brook, NY.
- Brower, L., Castilleja, G., Peralta, A., Lopez-Garcia, J., Bojorquez-Tapia, L., Díaz, S., et al. (2002). Quantitative changes in Forest quality in a principal overwintering area of the monarch butterfly in Mexico, 1971–1999. *Conserv. Biol.* 16, 346–359. doi: 10.1046/j.1523-1739.2002.00572.x
- Brower, L. P., Taylor, O. R., Williams, E. H., Slayback, D. A., Zubieta, R. R., and Ramirez, M. I. (2012). Decline of monarch butterflies overwintering in Mexico: is the migratory phenomenon at risk?: decline of monarch butterflies in Mexico. *Insect Conserv. Divers.* 5, 95–100. doi: 10.1111/j.1752-4598.2011.00142.x
- Coyle, E. J., Allebach, J. P., and Garton Krueger, J. (2006) The vertically-integrated projects (VIP) program in ECE at Purdue: fully integrating undergraduate education and graduate research, In Proceedings of the 2006 ASEE annual conference and exposition. 2006 ASEE annual conference and exposition, Chicago, IL.
- Ehlers, U. D. (2020) *Future skills: the future of learning and higher education*. Karlsruhe, Germany: Springer Books.
- European Commission, Executive Agency for Small and Medium Sized Enterprises. (2020) Skills for industry curriculum guidelines 4.0: future proof education and training for manufacturing in Europe: final report. LU: Publications Office. Available at: <https://data.europa.eu/doi/10.2826/097323> (Accessed June 19, 2023).
- Fernández, L. (2023). Educar en tiempos de cambio climático. Available at: <https://www.facebook.com/IFEconference/videos/ife-conference-talks-educar-en-tiempos-de-cambio-clim%C3%A1tico/234226172996420> (Accessed October 4, 2023).
- Frisk, E., and Larson, K. (2011). Educating for sustainability: competencies & practices for transformative action. *J. Sustain. Educ.* 2, 1–20.
- García-García, J. A., Reding-Bernal, A., and López-Alvarenga, J. C. (2013). Cálculo del tamaño de la muestra en investigación en educación médica. *Investig. Educ. Médica* 2, 217–224. doi: 10.1016/S2007-5057(13)72715-7
- Gómez Puente, S. M., and Doulougeri, K. (2022). “Challenge-based learning curriculum development: a suitable framework for engineering education” in *Towards a new future in engineering education, new scenarios that European alliances of tech universities open up* (Catalunya: Universitat Politècnica de Catalunya), 1196–1205.
- Kopnina, H. (2020). Education for the future? Critical evaluation of education for sustainable development goals. *J. Environ. Educ.* 51, 280–291. doi: 10.1080/00958964.2019.1710444
- López-Gómez, E. (2015). Conectando investigación y docencia en la Universidad: Teaching research nexus. *Revista Interuniversitaria*. 22, 203–220.
- Latinoamerica Verde (2018) Top 500 (Finalistas y Ganadores) - Premios verdes. premios latinoamérica verde. Available at: <https://premiosverdes.org/es/top-500-finalistas-y-ganadores/>.
- Mazur, B., and Walczynna, A. (2022). Sustainable development competences of engineering students in light of the industry 5.0 concept. *Sustainability* 14:7233. doi: 10.3390/su14127233
- Morini, C., Azevedo, A. T., Inacio Junior, E., and Coyle, E. J. (2022). A vertically integrated project (VIP) as a learning experience to foster entrepreneurship education. *Revista Empreendedorismo Gestão Pequenas Empresas*:e2333. doi: 10.14211/ibjesb.e2333
- Nehme, M. (2012). The nexus between teaching and research – easier said than done. *Leg. Educ. Rev.* 22, 241–272. Available at: <https://ssrn.com/abstract=2350571>
- Peña-Becerril, M., and Camacho-Zuñiga, C. (2020). Context as a strategy to promote the perceived usefulness of mathematics in social science students. *Formación Universitaria* 13, 145–156. doi: 10.4067/S0718-50062020000100145
- Pérez, J. M., Furman, D. A., Alemany, L. A., and Luque, F. (2022) RoBERTuito: a pre-trained language model for social media text in Spanish, In Proceedings of the thirteenth language resources and evaluation conference. Marseille, France: European Language Resources Association, 7235–7243
- Premios Verdes Team (2023) *Find out everything about Premios Verdes*. Available at: <https://premiosverdes.org/en/the-awards/> (Accessed June 20, 2023).
- Ryan, R. M., and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* 55, 68–78. doi: 10.1037//0003-066x.55.1.68
- Schaper, M. (2016) *Making ecopreneurs: Developing sustainable entrepreneurship*. 2nd ed. London: Routledge
- Sinclair, S., and Rockwell, G. (2016) Voyant tools. Available at: <http://voyant-tools.org/>
- Sonnenberg-Klein, J., and Coyle, E. J. (2022) Experience vs. academic rank: leadership growth over multiple semesters in vertically integrated project (VIP) courses, In 2022 IEEE Frontiers in education conference (FIE). 2022 IEEE Frontiers in education conference (FIE), Uppsala, Sweden: IEEE, 1–5

UNESCO. (2015) Sustainable development goals, UNESCO. Available at: <https://en.unesco.org/sustainabledevelopmentgoals> (Accessed October 18, 2022).

UNESCO and UNESCO Global Independent Expert Group on the Universities and the 2030 Agenda (2022). *Knowledge-driven actions: transforming higher education for global sustainability*. Paris, France: UNESCO.

United Nations (2023) *Global sustainable development report 2023*. Available at: <https://sdgs.un.org/gsdrgsdr2023> (Accessed June 20, 2023).

Wamsler, C. (2020). Education for sustainability: fostering a more conscious society and transformation towards sustainability. *Int. J. Sustain. High. Educ.* 21, 112–130. doi: 10.1108/IJSHE-04-2019-0152



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Mauro Mocerino,
Curtin University, Australia
Kfukfu Nsangong,
St. Cloud State University, United States

*CORRESPONDENCE

Yocanxóchitl Perfecto-Avalos
✉ yocan@tec.mx

RECEIVED 02 September 2023

ACCEPTED 25 January 2024

PUBLISHED 13 February 2024

CITATION

Sánchez-López AL, Jáuregui-Jáuregui JA,
García-Carrera NA and
Perfecto-Avalos Y (2024) Evaluating
effectiveness of immersive virtual reality in
promoting students' learning and
engagement: a case study of analytical
biotechnology engineering course.
Front. Educ. 9:1287615.
doi: 10.3389/feduc.2024.1287615

COPYRIGHT

© 2024 Sánchez-López, Jáuregui-Jáuregui,
García-Carrera and Perfecto-Avalos. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

Evaluating effectiveness of immersive virtual reality in promoting students' learning and engagement: a case study of analytical biotechnology engineering course

Angélica Lizeth Sánchez-López,
Jesús Antonio Jáuregui-Jáuregui, Nancy Anabel García-Carrera
and Yocanxóchitl Perfecto-Avalos*

Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Zapopan, Mexico

We conducted an immersive virtual reality (IVR) intervention in an analytical biotechnology course to evaluate its effectiveness in promoting student learning and engagement. The objective was to assess the impact of an IVR tool on learning infrared (IR) spectroscopy and academic engagement. The IVR experience was integrated into the course sessions, which included theoretical framework, IR spectra exercises, an individual quiz, and competition games. Students were immersed in a first-person IVR experience simulating an analytical room, where they performed sample analysis and operated a virtual IR spectrometer. The quiz and survey assessed their understanding and perception of the IVR experience. When compared to control groups, IVR intervention group achieved higher scores on midterm exam questions related to IR spectroscopy and reported greater academic engagement according to the satisfaction survey. The findings highlighted the effectiveness of IVR in enhancing learning outcomes, curiosity, motivation, and engagement among students. Moreover, our research provides empirical evidence of IVR's positive impact on student learning and engagement in the context of biotechnology engineering. Based on the study's implications, educators in engineering fields should consider integrating IVR as an instructional tool, particularly for laboratory-related topics with limited equipment accessibility and overcrowded courses. Future research should explore the generalizability of these findings across different subjects and educational levels. Additionally, investigating the underlying factors influencing the effectiveness of IVR in promoting academic engagement would further enhance the understanding and implementation of this technology in educational settings.

KEYWORDS

immersive virtual reality, academic engagement, experiential learning, infrared spectroscopy, biotechnology engineering, educational innovation, higher education, professional education

1 Introduction

Nowadays, the context of higher education demands educational strategies based on technological tools that facilitate the understanding of complex concepts, developing problem-solving skills and at the same time increasing the interest and motivation of students. Immersive learning is an emerging technology in the field of education, which allows the student to participate in the learning process by achieving meaningful and engaging interactions.

Virtual reality (VR) is implemented through two modalities based on the user experience. From one side, there is the modality based on the use of a desktop computer, in which the user has access to the content through a screen and interacts with it through the keyboard and mouse (Lee, 2014). The second modality, as described by Freina and Ott (2015), utilizes virtual reality headsets (VRH) as a form of immersive virtual reality (IVR) experience. This mode is characterized using a device that shows the user different views of a visual content as he turns his head. In addition, the user can complement his experience with handheld hearing aids and control devices.

Commonly, it is recognized that the use of VR in education has the advantages that it stimulates motivation to learn, improves attention to the object of study and allows experiential learning (Lei et al., 2018). In the case of education in the engineering field, the use of VR has allowed to simulate in real time the use of laboratory equipment that is difficult to access, it prevents students from causing damage to laboratory equipment, it solves the problem of lack of direct contact with laboratory equipment in overcrowded courses, reduces occupational risks and allows students to interact with complete processes (Vergara et al., 2017).

Structural chemistry is an important branch of chemistry focused on the study of the spatial arrangements of atoms and molecules, as well as the relationship between structure and properties (Zhou, 1993). Various IVR tools have been developed for learning chemistry, ranging from the visualization and manipulation of molecules in an immersive virtual environment to the simulation of practical experiences laboratories, either as organic chemistry lab (Tee et al., 2018; Camel et al., 2020; Williams et al., 2022), lab safety (Smith, 2016; Zhu et al., 2018; Unimersiv, 2019), molecular dynamics in enzyme catalysis (Bennie et al., 2019), and analytical instrumentation design (Naese et al., 2019). Regarding structural chemistry teaching, reports stand out because they test the visualization and manipulation of molecules in the immersive virtual environment, such as the construction of hydrocarbon molecules (Edwards et al., 2019), drug design (Norrby et al., 2015), the movement of molecules through carbon nanotubes (Ferrell et al., 2019), and the obtaining and interpretation of IR spectra to elucidate molecular structure (Dunnagan et al., 2020). In these studies, evidence is reported that the use of IVR generates interest, motivation, engagement, and greater learning.

Particularly, Dunnagan et al. (2020) developed and evaluated how an IVR tool could replace the use of equipment in an organic chemistry laboratory course. The authors designed a virtual immersive experience with a first-person perspective, comprised by the following stages:

1. An instructor greets the student, provides safety instructions, and offers an overview of the laboratory.
2. The student is taken to an analytical room where the IR spectrometer, a computer screen, four samples of unknown

compounds and cleaning material are located. Here, the student receives the necessary explanations for obtaining the IR spectrum of one of the samples. Finally, the student carries out the routine of analyzing the sample and cleaning the virtual machine.

3. The student leaves the analytical room, and the IR spectrum of the unknown compound is presented on a blackboard. During the analysis, the student must select sections of the spectrum and identify the frequencies to determine the corresponding functional groups and finally identify the unknown compound.

To assess the experience, the authors divided the students into “treatment” and “control” groups. At the end of the practical activity, students from both groups completed identical questionnaires, and they took the same exam at the end of the semester. The comparison of responses indicated that, although the use of IVR is more memorable, there is no statistically significant difference between the two learning methods. However, the results also suggested that the students who had the IVR experience retained learning better than those who had the conventional experience.

In recent decades, researchers and educators have shown a growing interest in the concept of engagement as a way to avoid student boredom, improve student motivation, and participate in school-related activities (Baron and Corbin, 2012), and increase the levels of student performance (Phan and Ngu, 2014), where the intrinsic motivation for studies called “academic engagement” is an important element to take into account.

Engagement has been recognized as a complex term that encompasses various patterns of motivation, cognition, and behavior. It occurs when students invest themselves psychologically in their learning, striving to understand the materials presented to them and internalize them into their lives (Appleton et al., 2008). This characteristic has been widely recognized as an important influence on achievement and learning in higher education (Kahu, 2013) and can be employed as a significant predictor of academic performance (Upadyaya and Salmela-Aro, 2013; Lee, 2014). Furthermore, engagement has been described that it improves the participation and motivation of students in STEM courses (science, technology, engineering and mathematics) in general, and specifically in organic and structural chemistry (Rossi, 2015).

The course “Analytical Aspects in Bioprocesses” is a subject in the Biotechnology Engineering (IBT) major that provides students with a fundamental understanding of instrumental analysis. One of the topics covered in this course is infrared (IR) spectroscopy, which incorporates concepts of molecular structure, chemical bonding, functional groups, and electromagnetic radiation. Traditionally, learning content related to analytical chemistry is supported by practical activities in a laboratory. However, “Analytical Aspects in Bioprocesses” is a theoretical course that does not involve laboratory activities, making the subject more abstract for most students. To address this challenge, we proposed the use of IVR in the study of IR spectra interpretation to improve the learning experience and increase academic engagement, while solving the lack of costly instrumentation and delicate handling by students.

This study aims to evaluate the use of an IVR tool as part of an integral didactic sequence for learning IR spectra interpretation in the course “Analytical Aspects in Bioprocesses.” We assessed the students’ learning by means of an individual midterm exam and compared the

scores with those obtained by students that did not use the IVR tool. Also, we evaluated students' attitudes and opinions regarding the use of IVR to complement the learning of IR spectra interpretation.

2 Materials and methods

The didactic sequence proposed in this study was carried out in the “Analytical Aspects in Bioprocesses” course during the February–June 2020 semester (FJ20), preceding the COVID-19 pandemic quarantine. This course is typically undertaken by IBT students in the eighth semester.

The didactic sequence, as depicted in Figure 1, unfolded through various steps. The initial session focused on the theoretical framework of IR, coupled with practical exercises using IR spectra in identifying organic compounds. In a second session, the students had the IVR experience proposed by Dunnagan et al. After the IVR experience, the students took an individual quiz, and the lecturer conducted a survey to collect the students' attitudes and opinions about the IVR experience. In the third and last session, students were asked to make randomized teams and participate in a competitive team game activity. It is noteworthy that while the IVR activity was a central component, the didactic sequence was designed to be integrative, complementing Dunnagan's IVR tool with additional pedagogical elements.

The assessment tools used during the didactic sequence are described next. The individual quiz required students to associate an IR spectrum with five compounds identified by their chemical structure. The survey, administered anonymously via Google Forms, encompassed two sections. The first one collected general demographic data, such as gender, age, semester, minor, and origin of VR lenses. The second section comprised 12 questions assessing the student's perception of the IVR experience, measured using a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). This scale is a widely used measurement tool in educational research and is designed to gather information about individuals' attitudes, opinions, and behaviors by presenting a series of statements and asking respondents to indicate their level of agreement or disagreement (Kusmaryono and Wijayanti, 2022). To maintain the validity of the questionnaire, we ensured that the statements were clear, unbiased and cover the entire range of possible responses. Lastly,

during the competitive team activity, students were tasked with associating 10 IR spectra with their respective compounds, identifying them by both their chemical structure and generic name.

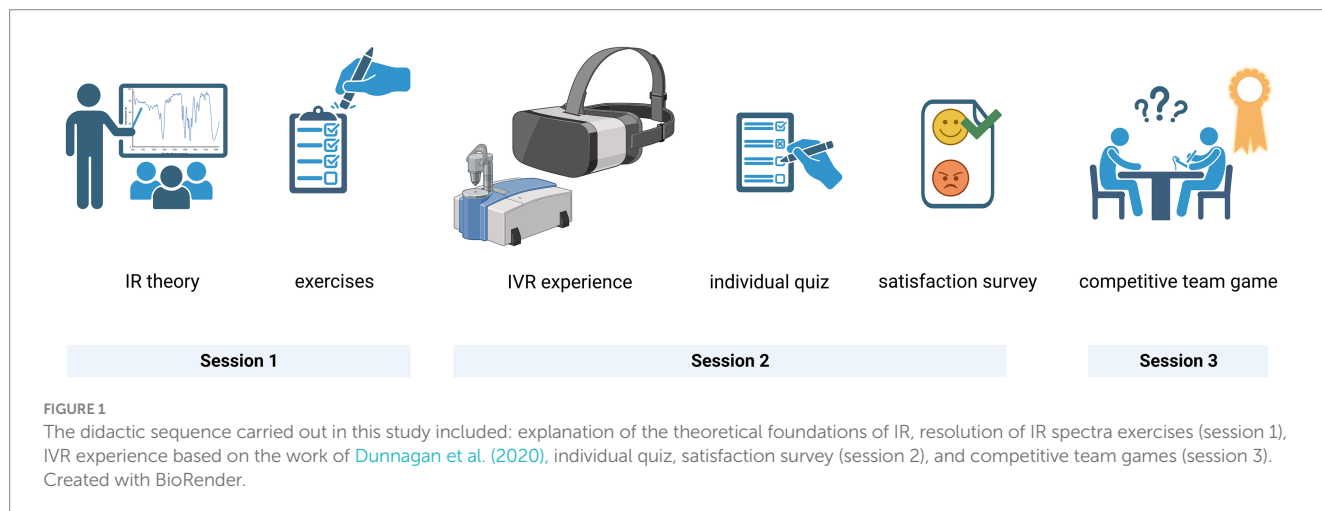
The students also took a midterm individual exam including a section with IR questions. The scores from these questions were retrieved to be compared with the corresponding ones in historical results from the January–May 2017 (JM17), 2018 (JM18), and 2019 (JM19) groups. It is important to clarify that the didactic sequence for JM17, JM18, and JM19 groups included only the activities described for sessions 1 and 3.

Quantitative data from the individual quiz, survey, team game, and the midterm individual exam were drawn and analyzed. Statistical analysis employed the use of difference in proportions for survey results and analysis of variance for comparing group and team performance in the game and the midterm individual exam, respectively. A significance level of p -value < 0.05 was considered, and data processing was carried out with Minitab (v19, Minitab, Inc., United States) and RStudio (v1.3.1056, RStudio, United States). In addition to collecting quantitative data, the lecturer conducted a qualitative observation of students' mood during the IVR experience. This qualitative aspect, solely for descriptive purposes, was not quantified or subjected to formal analysis.

3 Results

The FJ20 group, in which the IVR experience was implemented, consisted of 39 students, with 21 women (55.3%) and 17 men (44.7%) (Figure 2A). The students had a mean age of 22 years old (± 0.9 years). In terms of minors (specialization area), 31 students (79.5%) were from Bioprocesses (BPR), and 8 students (20.5%) were from Molecular Biology (MOL). As depicted in Figure 2B, for the IVR experience, approximately half of the students borrowed VR headsets (54%), one-third of the group purchased them (31%), some students had prior experience with VR (10%), and a few students had the creativity to make their own headsets (5%).

The IVR experience took place in the classroom and lasted approximately 30 min. The students were relaxed and actively engaged during the session. At the end of the IVR experience, the students answered an individual quiz, obtaining an average score of 98.4%



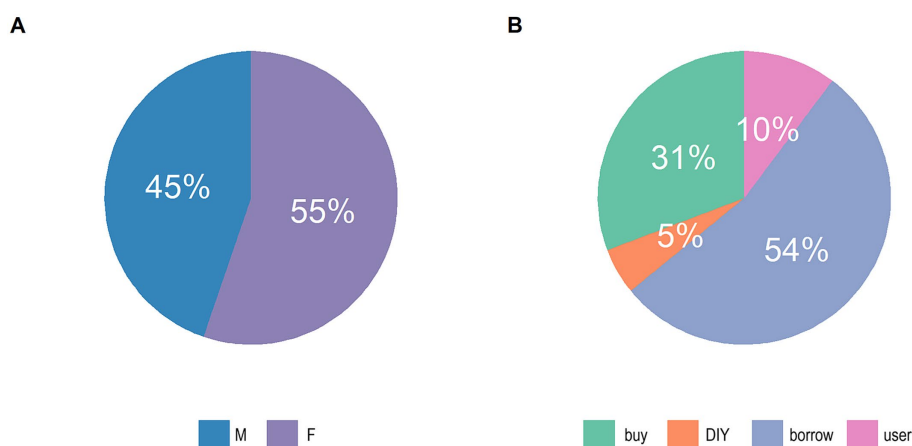


FIGURE 2
(A) IBT students' characteristics. (B) VR headset source. DIY, Do it yourself.

($\pm 5.5\%$). This suggests a consistent understanding of the theoretical aspects covered both in class and with the IVR experience. Following the quiz, the students were asked to participate in a survey, and all students responded.

The students had a highly positive perception of the IVR experience, as indicated by the percentage of students who agreed or strongly agreed on the Likert scale (4 and 5, respectively). As shown in Figure 3, the students considered the experience to be immersive (84.6%), with learning activities that related to real-life situations (92.3%). They also reported enjoying the activities (89.7%) and expressed interest in repeating the experience to learn other analytical techniques (84.6%). The IVR experience was deemed helpful in understanding the theory effectively (79.5%), facilitating the completion of the individual quiz (76.9%), and most students (92.3%) would recommend IVR technology as a learning tool. Notably, in response to the question regarding recommendation of IVR technology, a higher proportion of women indicated agreement compared to men ($p=0.045$). These results suggest that the IVR experience was very appealing and satisfying to the students, in addition to facilitating the understanding of the content.

The students of FJ20 and historic groups JM17, JM18, and JM19 played in teams in the association game of 10 IR spectra with 10 molecules. This occurred in their respective courses. As shown in Figure 4A, a significant difference was not found between the score of the group that experienced IVR (FJ20) and the results from previous academic periods. Regarding the FJ20 game performance, there was also no significant difference between the teams, and only 2 teams had less than 6 points (teams D4 and D9, Figure 4B). When comparing the IR questions score included in the midterm exam of the FJ20 group, by the teams organized during the game, we did not find any difference either.

It is important to mention that teams with the highest scores in the game also demonstrated the highest scores in the IR questions of the midterm exam (teams D3 and D5, Figure 4B). Out of the 39 students, only 2 displayed poor performance in the IR questions of the midterm exam. Similar data could not be retrieved from previous periods. In this sense, our findings align with those reported by the developers of the IVR tool used (Dunnagan et al., 2020), who found no statistically significant difference between the use of IVR and

conventional learning. However, we believe that by complementing the theoretical contents with the IVR experience, an otherwise arid topic became interesting and attractive to students. 4,175.

4 Discussion

There are still a limited number of IVR resources available for teaching topics such as instrumental analysis. Particularly, Naese et al. (2019) used augmented reality technology to teach the design and operation of four analytical instruments: flame atomic absorption spectrometer, gas chromatograph–mass spectrometer, liquid chromatograph, and double-beam UV–vis spectrophotometer. It was noted that, although the participating students liked using augmented reality to learn more about the instruments used in class, they had mixed views when prompted to further comment on the activity. This was mainly because of difficulties with the app when accessing the displayed information when not in front of the instrument. Moreover, the main limitation of this specific augmented reality (AR) tool is that it cannot be deployed in other universities unless the exact equipment is available. However, it is possible that the same or a similar AR platform could be used to visualize the components and operation of other types of equipment, but this would likely require the creation of new content specific to the equipment in question.

Conversely, Dunnagan et al. tool was based on recorded sessions. The VR experience was designed to be an immersive experience that would be as effective as the being in the lab. It comprised both teaching students how to operate an IR spectrometer, and how to elucidate simple features of an IR spectrum. This was the experience we needed for the infrared spectroscopy lecture, particularly because the course was a theoretical one, and the lack of this equipment in our labs.

In a second study, Dunnagan reported that the undergraduate students who tried the IVR experience during the pandemic reported a high degree of satisfaction with the tool and no significant usability barrier (Dunnagan and Gallardo-Williams, 2020). Nonetheless, some of the students were not comfortable with the experience, given that they found their plans for a face-to-face lab experience disrupted by the unexpected quarantine situation. In contrast, we deployed the IVR tool weeks before the quarantine, and our results shows a highly



FIGURE 3
Survey results after IVR experience.

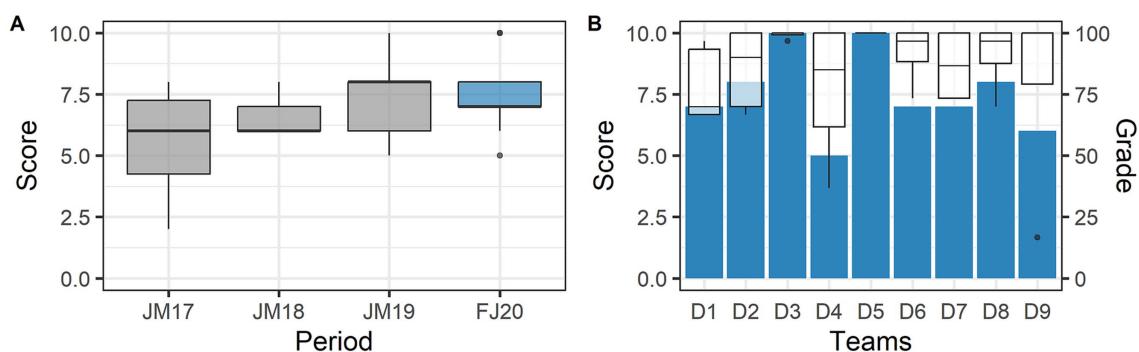


FIGURE 4
(A) Score obtained in IR game of groups JM17, JM18, JM19, and FJ20. (B) Performance comparative by teams in IR game (blue bars) and questions scores (boxplots) included in midterm exam of group FJ20.

appreciated learning experience, related to real-life situations. Also, our students showed a keen interest in repeating the experience to learn other analytical techniques, and most of them would recommend IVR technology as a learning tool. It is noteworthy that the students found creative ways to overcome the lack of VR headsets, which demonstrates their keen interest in participating in the IVR experience. The student's positive perception to augmented and/or immersive virtual reality tools have been reported in numerous other topics. Students have benefited from the VR technology to aid the learning experience. Bennie et al. (2019) found that while learning

about molecular dynamics in enzyme catalysis, undergraduate students found the VR platform enjoyable and easy to use. The authors did not attempt to quantify the direct learning gains of the participants, but students that experienced VR tool in enzyme catalysis obtained a higher median score than the ones who did not used it.

In light of our objectives, this study demonstrated significant positive outcomes. We evaluated the effectiveness of using an IVR experience as a support tool for learning an analytical technique. Although we did not observe a significant improvement in student learning compared to other courses, the results of the satisfaction

survey indicated that the IVR experience was highly valued by the students, and it successfully enhanced their engagement. The findings also underscore the broader potential of IVR in reducing difficulties related to abstract course content, limited access to specialized equipment, and large class sizes.

5 Limitations and future research

While this study's results highlight how IVR might improve student learning and engagement, certain limitations merit consideration. First off, the study's applicability was restricted to a particular course, a specific analytical technique, and a particular instrument, so it's unclear whether the findings can be applied in other educational settings or to another analytical methods. Second, the evaluation focused on the short term results of the IVR intervention, mostly by means of a satisfaction survey and midterm exam that were conducted soon after the IVR experience. Third, it would be interesting how individual learning styles, backgrounds, or past experiences with virtual reality technology can affect students' reactions and affect how successful is the IVR tool.

Addressing the identified limitations opens possibilities for future research to enrich our understanding of IVR tools in educational settings. For example, exploring the tool's efficacy in a variety of academic fields and classes would offer a more comprehensive view of its versatility and impact. Also, a few weeks or months following the IVR intervention follow-up evaluations may be carried out to see how long the advantages last and to provide guidance on how to incorporate IVR into curriculum. A study comprising more students in remote and in-person courses, from different campus, different majors, and institutions, would shed light on this matter. Also, the transferability of IVR strategy to the use of real equipment will solve training concerns for costly equipment.

In conclusion, the use of IVR proved to be an exceptional supportive feature for a course that lacked hands-on lab activities. The high level of engagement generated through the IVR experience suggests that this tool has the potential to be considered not only in traditional classroom settings but also in remote learning environments.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

References

- Appleton, J. J., Christenson, S. L., and Furlong, M. J. (2008). Student engagement with school: critical conceptual and methodological issues of the construct. *Psychol. Sch.* 45, 369–386. doi: 10.1002/pits.20303
- Baron, P., and Corbin, L. (2012). Student engagement: rhetoric and reality. *High. Educ. Res. Dev.* 31, 759–772. doi: 10.1080/07294360.2012.655711
- Bennie, S. J., Ranaghan, K. E., Deeks, H., Goldsmith, H. E., O'Connor, M. B., Mulholland, A. J., et al. (2019). Teaching enzyme catalysis using interactive molecular dynamics in virtual reality. *J. Chem. Educ.* 96, 2488–2496. doi: 10.1021/acs.jchemed.9b00181
- Camel, V., Maillard, M.-N., Piard, J., Dumas, C., Cladière, M., Fitoussi, G., et al. (2020). CHIMACTIV: an open-access website for student-centered learning in analytical chemistry. *J. Chem. Educ.* 97, 2319–2326. doi: 10.1021/acs.jchemed.0c00023
- Dunnagan, C. L., Dannenberg, D. A., Cuares, M. P., Earnest, A. D., Gurnsey, R. M., and Gallardo-Williams, M. T. (2020). Production and evaluation of a realistic immersive virtual reality organic chemistry laboratory experience: infrared spectroscopy. *J. Chem. Educ.* 97, 258–262. doi: 10.1021/acs.jchemed.9b00705
- Dunnagan, C. L., and Gallardo-Williams, M. T. (2020). Overcoming physical separation during COVID-19 using virtual reality in organic chemistry laboratories. *J. Chem. Educ.* 97, 3060–3063. doi: 10.1021/acs.jchemed.0c00548
- Edwards, B. I., Bielawski, K. S., Prada, R., and Cheok, A. D. (2019). Haptic virtual reality and immersive learning for enhanced organic chemistry instruction. *Virtual Real.* 23, 363–373. doi: 10.1007/s10055-018-0345-4
- Ferrell, J. B., Campbell, J. P., McCarthy, D. R., McKay, K. T., Hensinger, M., Srinivasan, R., et al. (2019). Chemical exploration with virtual reality in organic teaching laboratories. *J. Chem. Educ.* 96, 1961–1966. doi: 10.1021/acs.jchemed.9b00036

Author contributions

AS-L: Investigation, Methodology, Writing – original draft, Writing – review & editing. JJ-J: Investigation, Writing – original draft, Methodology, Writing – review & editing. NG-C: Writing – original draft. YP-A: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Acknowledgments

The authors would like to express their gratitude to the Department of Bioengineering at the Guadalajara campus for providing the necessary facilities to implement this project. In addition, the authors acknowledge the Emerging Food Technologies and Nutraceuticals group. The authors would like to acknowledge the financial support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico, in the production of this work.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Freina, L., and Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives. In *Rethinking education by leveraging the eLearning pillar of the digital agenda for Europe* (Carol I National Defence University Publishing House), 133–141.
- Kahu, E. R. (2013). Framing student engagement in higher education. *Stud. High. Educ.* 38, 758–773. doi: 10.1080/03075079.2011.598505
- Kusmaryono, I., and Wijayanti, D. (2022). Number of response options, reliability, validity, and potential Bias in the use of the Likert scale education and social science research: a literature review. *Int. J. Educ. Methodol.* 8, 625–637. doi: 10.12973/ijem.8.4.625
- Lee, J.-S. (2014). The relationship between student engagement and academic performance: is it a myth or reality? *J. Educ. Res.* 107, 177–185. doi: 10.1080/00220671.2013.807491
- Lei, X., Zhang, A., Wang, B., and Rau, P.-L. P. (2018). “Can virtual reality help children learn mathematics better? The application of VR headset in Children’s discipline education” in *Cross-cultural design. Applications in cultural heritage, creativity and social development, Lecture notes in computer science*. ed. P.-L. P. Rau (Cham: Springer International Publishing), 60–69.
- Naese, J. A., McAteer, D., Hughes, K. D., Kelbon, C., Mugweru, A., and Grinias, J. P. (2019). Use of augmented reality in the instruction of analytical instrumentation design. *J. Chem. Educ.* 96, 593–596. doi: 10.1021/acs.jchemed.8b00794
- Norrby, M., Grebner, C., Eriksson, J., and Boström, J. (2015). Molecular rift: virtual reality for drug designers. *J. Chem. Inf. Model.* 55, 2475–2484. doi: 10.1021/acs.jcim.5b00544
- Phan, H. P., and Ngu, B. H. (2014). An empirical analysis of students’ learning and achievements: a motivational approach. *Educ. J.* 3:203. doi: 10.11648/j.edu.20140304.11
- Rossi, R. D. (2015). ConfChem conference on flipped classroom: improving student engagement in organic chemistry using the inverted classroom model. *J. Chem. Educ.* 92, 1577–1579. doi: 10.1021/ed500899e
- Smith, C. (2016). Chemistry lab VR. *Dev. Dent.* Available at: <https://devpost.com/software/chemistry-lab-vr> (Accessed June 24, 2023).
- Tee, N. Y. K., Gan, H. S., Li, J., Cheong, B. H.-P., Tan, H. Y., Liew, O. W., et al. (2018). Developing and demonstrating an augmented reality colorimetric titration tool. *J. Chem. Educ.* 95, 393–399. doi: 10.1021/acs.jchemed.7b00618
- Unimersiv (2019). Chemistry VR. *Unimersiv*. Available at: <https://unimersiv.com/review/chemistry-vr/> (Accessed June 25, 2023).
- Upadyaya, K., and Salmela-Aro, K. (2013). Development of school engagement in association with academic success and well-being in varying social contexts. *Eur. Psychol.* 18, 136–147. doi: 10.1027/1016-9040/a000143
- Vergara, D., Rubio, M. P., and Lorenzo, M. (2017). On the design of virtual reality learning environments in engineering. *Multimodal Technol. Interact.* 1:11. doi: 10.3390/mti1020011
- Williams, N. D., Gallardo-Williams, M. T., Griffith, E. H., and Bretz, S. L. (2022). Investigating meaningful learning in virtual reality organic chemistry laboratories. *J. Chem. Educ.* 99, 1100–1105. doi: 10.1021/acs.jchemed.1c00476
- Zhou, G. (1993). *Fundamentals of structural chemistry*. Singapore, World Scientific.
- Zhu, B., Feng, M., Lowe, H., Kesselman, J., Harrison, L., and Dempsey, R. E. (2018). Increasing enthusiasm and enhancing learning for biochemistry-laboratory safety with an augmented-reality program. *J. Chem. Educ.* 95, 1747–1754. doi: 10.1021/acs.jchemed.8b00116



OPEN ACCESS

EDITED BY

Julius Nganji,
University of Toronto, Canada

REVIEWED BY

Marios A. Pappas,
National Centre of Scientific Research
Demokritos, Greece
Asma Khattala,
University of Sétif 2, Algeria

*CORRESPONDENCE

Claudia Camacho-Zuñiga
✉ claudia.camacho@tec.mx

RECEIVED 30 June 2023

ACCEPTED 28 February 2024

PUBLISHED 13 March 2024

CITATION

Miralrio A, Muñoz-Villota J and
Camacho-Zuñiga C (2024) From flexibility to
adaptive learning: a pre-COVID-19
perspective on distance education in Latin
America.

Front. Comput. Sci. 6:1250992.

doi: 10.3389/fcomp.2024.1250992

COPYRIGHT

© 2024 Miralrio, Muñoz-Villota and
Camacho-Zuñiga. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

From flexibility to adaptive learning: a pre-COVID-19 perspective on distance education in Latin America

Alan Miralrio¹, Jessica Muñoz-Villota² and
Claudia Camacho-Zuñiga^{1,3*}

¹Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Monterrey, Mexico, ²Universidad Pedagógica Nacional, Unidad Ajusco, Mexico City, Mexico, ³Institute for the Future of Education, Tecnológico de Monterrey, Monterrey, Mexico

At the time of this writing, emerging Generative Artificial Intelligence (GenAI) technologies, such as ChatGPT, Bard, and DALL-E, are shaping education. Despite their great potential for enhancing education, the discussion on the risks for the Society is an ongoing debate. Remote or distance education (DE) in developed countries has evolved in parallel to the permeation of Information and Communication Technologies (ICT). In Latin America (LATAM), a deep understanding of this interrelated evolution will encourage an efficient implementation of innovative policies, pedagogies, and technologies, including GenAI. This paper presents the historical development of DE in LATAM, drawing connections with milestones along the evolution of ICT in the region. This evolution is described across five generations: Correspondence, Audiovisual, ICTs-based, Web-based, and Interactive. Each generation offered incremental benefits to students, from flexibility and well-designed instructional materials to adaptive, interactive, and cost-effective education. This research encompassed a comprehensive search on Scopus, Web of Science, Google Scholar, ScienceDirect, Latindex, Dialnet, Redalyc, and SciELO with an historical approach. It yielded 97 peer-reviewed articles, books, and reports from recognized international organizations, published in Spanish or English and covering a period spanning from the end of the 19th century to the onset of COVID-19 lockdown. The analysis confirms the critical role of Higher Education Institutions (HEIs) in LATAM in supporting ICT integration, demonstrating a synergistic interplay with DE. Notwithstanding, the region's social inequality and digital divide have delayed the full deployment of DE's advantages, in comparison to USA and Europe. The paper exposes different scenarios and tech-educational requirements of DE, showcasing Mexico, Brazil, and Colombia, as well as other countries with smaller populations. Notably, by 2019, *Tecnológico de Monterrey* in Mexico, and *Universidade de São Paulo* and *Universidade Federal da Paraíba* in Brazil have pioneered the implementation of the Fifth Generation of DE in the region. The findings of this research highlight the value of DE as a cost-effective solution to traditional education, addressing social barriers and broadening educational access in LATAM, especially during emergencies. This research poses significant implications that can equip digital technology providers, HEIs, governments, and policymakers as they navigate the ongoing educational challenges.

KEYWORDS

distance education, Latin America, remote education, information and communication technologies, distance learning, higher education, educational innovation

1 Introduction

The COVID-19 pandemic created an adverse scenario for every nation around the World, including severe impacts on the educational sectors. The [World Bank \(2020\)](#) synthesized these affectations in two points: the economic recession sparked by pandemic-control measures and the near-universal closure of schools at all levels.

The COVID-19 pandemic profoundly disrupted education worldwide, prompting an unprecedented shift to remote or distance education (DE) and driving our attention toward policies and strategies for Emergency Remote Teaching (ERT). UNESCO defined in 1993, as cited in [Moreno and Maluche \(2017\)](#) that DE is the “education that is based on non-face-to-face communication, meaning that during the educational process the student is at a certain distance from the teacher, either for part, most, or even the entire duration of the process, depending on different circumstances.” Similar definitions have been stated underlining the following characteristics of DE: A formal education supported by an institution; the physical separation between the instructor and the student; different means of communication to connect students, learning resources, and instructors as well as a flexible learning ([García Aretio, 1999; Simonson, 2006; Cabral, 2022](#)). Flexibility in DE is understood in terms of, but not limited to, the timing and location of learning activities, the pace at which students progress through the curriculum, the choice of learning materials, the class size—massive or not, and the interaction available to learners. Throughout this work, DE will be assumed as formal education modalities with the mentioned characteristics.

ERT stems in DE, since it enables delivering education even under crises, such as pandemics, natural disasters, wars, or other circumstances that hinder traditional, face-to-face educational activities. ERT relies on providing a resilient, adaptable, and accessible educational system to ensure the delivery of high-quality education during disruptions to traditional inside-facilities teaching and learning environments ([Crompton et al., 2022](#)).

Implementing ERT has been particularly challenging in Latin America (LATAM), where access to Information and Communication Technologies (ICT) is limited. As countries like China and Japan rapidly deployed online courses ([MEXT, 2020; UNESCO, 2020](#)), LATAM responses were more conservative, reflecting the region’s unique challenges. For instance, Mexico adopted a television-based approach to deliver basic education, broadcasting recorded lessons to overcome connectivity issues.

Inequalities in LATAM limited the implementation and broadening of DE among the population, restricting the possibilities to provide quality education during COVID-19 pandemic. In general, Latin-American higher education institutions (HEI) and governments took several years to join the Network; and the general population has had limited access to ICT ([Rama, 2016](#)). In fact, since 2014, the *Organización de Estados Iberoamericanos* warned about the impact of digital divide on social and educative tendencies in LATAM.

The digital divide highlights the relevance of understanding the historical development of DE and ICT in LATAM. This knowledge can provide valuable insights for policymakers, educators, and researchers as they navigate the current challenges and plan for the future. Moreover, with the advent of AI technologies like ChatGPT and Bard, which are increasingly pervading the educational sector, understanding the historical context becomes even more critical.

These technologies hold great potential for enhancing DE, but their successful implementation will depend on a nuanced understanding of the region’s educational history and infrastructure.

This paper provides a historical review of DE and ICT in Latin America, focusing on the development of policies, technological infrastructure, and pedagogical approaches. We begin by examining the early development of DE prior to the advent of ICT, then explore the initial adoption of ICT by Higher Education Institutions (HEIs), and finally, discuss the integration of ICT with DE. The historical viewpoint recognizes five generations of DE in synergy with technological evolution: First Generation or Correspondence DE; Second Generation or Audiovisual DE; Third Generation or ICT-based DE; Fourth Generation or Web-based DE, and finally, Fifth Generation or Interactive DE. We conclude by analyzing the evolution of DE in some LATAM countries prior to the COVID-19 pandemic, to gain insights into the challenges and opportunities of DE in the region, as well as the different contexts and the required technological tools.

Our aim is to provide a grounded understanding of the capacities and infrastructure of educational institutions across Latin America at the onset of the COVID-19 crisis. This understanding will not only shed light on the current state of DE in the region but also inform the effective integration of emerging AI technologies in education. As such, this review is of critical relevance to today’s HEIs and will continue to be pertinent as we navigate the future of education in the Digital Age.

2 Methods

2.1 Procedures

This review aims to provide a historical perspective on the evolution of distance education (DE) in Latin America (LATAM), with a particular focus on the relationship with the adoption and evolution of Information and Communication Technologies (ICT). Following the route of previous historical literature reviews ([Chanock, 2011; Remenick, 2019; Georgiou, 2021](#)), this work assumes a “traditional” approach based on narrative. According to [Georgiou \(2021\)](#), the historical approach is more than a chronological description within a discipline. It not only situates the reader within an “historical context”; but also emphasizes a critic evaluation of a topic. Thus, it provides a historical perspective for engaging the body of the literature that is useful to identify and illustrate the issues of a given topic ([Georgiou, 2021](#)). For example, [Remenick \(2019\)](#), in her historical literature review, elucidated the development of nontraditional models of students during four different historical time periods.

The American Educational Research Association recognizes the historical approach together with theoretical, methodological, and integrative reviews as eligible for educational research ([Kennedy, 2007](#)). The historical narrative literature review is not necessarily enclosed within systematic or nonsystematic reviews. It employs a flexible approach to literature selection and analysis that aims for a qualitative understanding of the evolution of DE and ICT. In contrast, for example, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) ([Page et al., 2021](#)) is designed toward systematic reviews and meta-analyses that synthesize empirical

research findings. To ensure transparency in the reporting process, though, we have detailed the sources consulted, the rationale for their inclusion, and the methodological approach to construct the narrative. See Sections 2.2 and 2.3.

In terms of time, we focus on the period prior to the COVID-19 pandemic, paying special attention to the most populated nations in LATAM: Mexico, Brazil, and Colombia to evidence the different contexts and tech-educational requirements in the region. Also, Argentina, Peru, Venezuela, and Puerto Rico were selected to complement the discussion.

Our methodology follows the procedures proposed by [Petticrew and Roberts \(2008\)](#), which include defining the research questions, determining the types of studies, conducting a comprehensive literature search, screening the search results, appraising the included studies, synthesizing the studies, and assessing heterogeneity among the studies.

After being actively involved in the process of shifting modality from face-to-face to online modality and vice versa, we posed the following research questions (RQ) to guide this work:

- RQ1: What were the key developments in distance education prior to the advent of Information and Communication Technologies?
- RQ2: What strategies did Higher Education Institutions employ in the initial adoption of Information and Communication Technologies and the Internet?
- RQ3: At what point and through what processes did Information and Communication Technologies integrate with distance education?
- RQ4: How has the implementation of Information and Communication Technologies in Latin American Education progressed over time?
- RQ5: What was the status of distance education in the three most populated Latin American countries prior to the COVID-19 pandemic?
- RQ6: How did the integration of Information and Communication Technologies influence the evolution of distance education in Latin America prior to the COVID-19 pandemic?

2.2 Databases and search terms

To identify, assess, and synthesize relevant studies to answer our research questions we followed a similar procedure to that reported by [Lee et al. \(2019\)](#). It included searching selected terms in databases, using Boolean operators and snowball methods ([Lee et al., 2019](#)). We selected the following databases to ensure a broad coverage of both international and Latin American literature: Scopus, Web of Science, Google Scholar, ScienceDirect, Latindex, Dialnet, Redalyc, and SciELO. Since highly cited Latin American authors frequently publish in Spanish, the search was conducted in both English and Spanish using the keywords listed in [Table 1](#). With respect to the search query, up to four different keywords of [Table 1](#) were used to build additional strings through the Boolean operators AND as well as OR. All documents were retrieved through the library websites of Tecnológico de Monterrey and Universidad Pedagógica Nacional. The next subsection details the selection criteria used to curate the results of the searching procedure.

TABLE 1 Keywords used to carry out the literature search through search engines.

Focus of interest	Distance education and ICT concepts
<ul style="list-style-type: none">• Latin America, LATAM• Mexico• Brazil• Colombia• Argentina• Peru• Venezuela• Puerto Rico• Higher education institutions	<ul style="list-style-type: none">• Distance education, DE• Correspondence-based education• Web-based education• Multimedia education• Interactive education• Audio-visual education• Information and communications technology, ICT• Epistolary education• Net-based education• Tele-education• Educational broadcasting• Radio in education• TV in education• Online education

2.3 Selection criteria

Initially, the literature included in this review was limited to peer-reviewed papers and books of recognized editorials, published in Spanish or English from January 2015 up to April 2020, while the search was conducted in December 2019. This time frame was strategically chosen due to the significant advancements in ICT that directly impacted the modalities and reach of DE, especially with the rise of Massive Open Online Courses (MOOCs) and the increased accessibility to high-speed internet in remote areas. This period encapsulates a transformative phase in DE, marked by technological innovations and global events, just before the onset of the COVID-19 lockdown, which accelerated DE adoption. It is also noteworthy that many of the included papers were comprehensive reviews on the development of DE at both a global scale and specifically within the LATAM region.

Initially, this search resulted in 92 publications. After a careful analysis of the abstracts and a screening of the whole documents, we excluded publications with similar topics, comparable strategies or approaches, and same authorship, resulting in 76 publications. The documents' topics to be included, i.e., the inclusion and exclusion criteria, are summarized in [Table 2](#).

After exploration of the references in the selected publications and to obtain more recent statistics and official announcements, we consulted supranational organizations and governments' webpages and white papers. Therefore, 21 more publications were included for the sake of a complete overview of the historical development of DE in the global context. As a result, our review covers a period spanning from the end of the 19th century to the onset of COVID-19 lockdown, providing a rich historical context that underscores the evolution and significance of DE over time.

2.4 Analysis

After a detailed reading of the full texts, we identified patterns and themes related to the evolution of DE. We categorized them in two groups: those providing a global context and historical progression of

TABLE 2 Criteria to determine the studies included in this review on distance education (DE) in LATAM.

Inclusion criteria	Exclusion criteria
<ul style="list-style-type: none"> • Papers that reviewed the development of DE at a global scale. • Publications that report the development of DE in LATAM region. • Studies comparing, quantitatively, the enrolment in face-to-face programs with DE alternatives. • Publications that provide statistical data and correlations between DE and social indicators. 	<ul style="list-style-type: none"> • Empirical studies examining the impact of DE implementations on the learning-teaching process. • Reports of a certain implementation of DE. • Studies reporting similar approaches to other papers already analyzed and discussed.

DE, and those offering details on the adoption and current state of DE in LATAM. We synthesized the themes and compared across texts to build a comprehensive narrative of the general evolution of DE that simultaneously deepened on relevant experiences within this region, including those in the most populated countries in the region, i.e., Brazil, Mexico, and Colombia. We also build a coherent timeline of DE development and the integration of ICT within HEIs, to serve as a scaffold to our narrative. Since educational level is a key indicator of socio-economic development, this aimed to interweave the ICTs global trends with regional experiences in DE.

3 Results and discussion

To answer the research questions that were defined above, the contents of the selected articles are analyzed and discussed through the rest of the paper.

3.1 RQ1: what were the key developments in distance education (DE) prior to the advent of Information and Communication Technologies (ICT)?

Despite DE has evolved through five different generations (Castillo et al., 2017), only two generations emerged prior to the advent of ICT. After the advent of these technologies, DE had a rapid evolution linked to the ICT fast-evolving pace and its increasing pervasion in the different sectors of humanity. The timeline shown in Figure 1 illustrates an overview of the facts relevant to the DE history. The following paragraphs will answer RQ1 by correlating historical milestones with the development of the different generations of DE.

The first generation was based on printed media sent by postal services and was named Correspondence Education (Taylor, 1995; García Aretio, 1999). Although García Aretio marked the beginning of correspondence education at the end of the 19th century, currently exists certain controversy. Toro and Rama (2013) suggested the earliest implementations of DE could be the hieroglyphs found in priest schools from the ancient Egypt as well as the clay printed tablets sent the by the Mesopotamian king Ashurbanipal to educate his people.

The British University Correspondence College was one of the first HEIs to implement correspondence education (Kanwar and Daniel, 2010). This private institution provided courses for students enrolled in the University of London, which had no teaching functions and only conferred degrees after examinations (Kaye, 1981). Moreover, other well-established universities, as the Illinois Wesleyan University (1874), the Queen's University in Kingston, Canada (1889), and the University of Chicago, US (1891), initiated correspondence courses (Kanwar and Daniel, 2010). Also remarkable was the Centre National d'Enseignement à Distance, established by the French government in 1939, to provide correspondence education during World War II (Lecourt, 1988).

The second generation of DE was named Multi-media or Audio-visual Education; it integrated radio, telephone, television, video, and telegraph (García Aretio, 1999). It emerged from industrial inventions and global conflicts. Toward the middle of the 20th century, DE played a special role in the reconstruction of Europe. After the consecutive World Wars, audio-visual education pushed the formation of skilled workers for the re-emerging industry and the educational sector (Toro and Rama, 2013).

Audio-visual education reached a milestone through the creation of radio and television-based programs at the University of South Africa (UNISA) in 1946. UNISA became, in 1962, the first HEI in the World to deliver only DE programs (Yong et al., 2017). During the 1960s and 1970s, several DE institutions were created around the World. In 1963, *Centro Nacional de Enseñanza Media por Radio y Televisión*, a Spanish middle education institution, started to offer television- and radio-based courses. In 1969 in United Kingdom, the Open University was created as an autonomous and independent DE organization (Bell and Tight, 1993). Three years later, this entity inspired the creation of *Universidad Nacional de Educación a Distancia*, as well as a large list of other Spanish HEI (García Aretio, 1999). This progressive growth in Spain led, in 1977, to the association of DE institutions under the name *Asociación Nacional de Centros de Enseñanza a Distancia* (García Aretio, 1999).

Certainly, the contenders of the Cold War, US and USSR, pushed DE to increase their bases of skilled workers and educated people. Starting in 1920 from print-based educational systems, by the 1970s the USSR had enrolled around one million students in distance programs (Zawacki-Richter and Kourotchikina, 2012). As collected by García Aretio (1999), from data of Rumble and Keegan, in 1979 more than 2,200,000 university students enrolled in distance programs, although correspondence-based yet. In contrast, North American institutions took advantage of the most recent technologies. As Yong et al. (2017) synthesized, Wisconsin University offered telephone-based courses since 1965; while Stanford University created the Stanford Instructional Network Television in 1967. Even more, the National Technological University was created in 1985 to provide satellite-based courses for engineering degrees.

3.2 RQ2: what strategies did Higher Education Institutions (HEIs) employ in the initial adoption of Information and Communication Technologies (ICT) and the internet?

Close to the beginning of the 21st Century, digital higher education was born from previous DE experiences (Yong et al., 2017) and it was developed in parallel with ICT, crystalized in

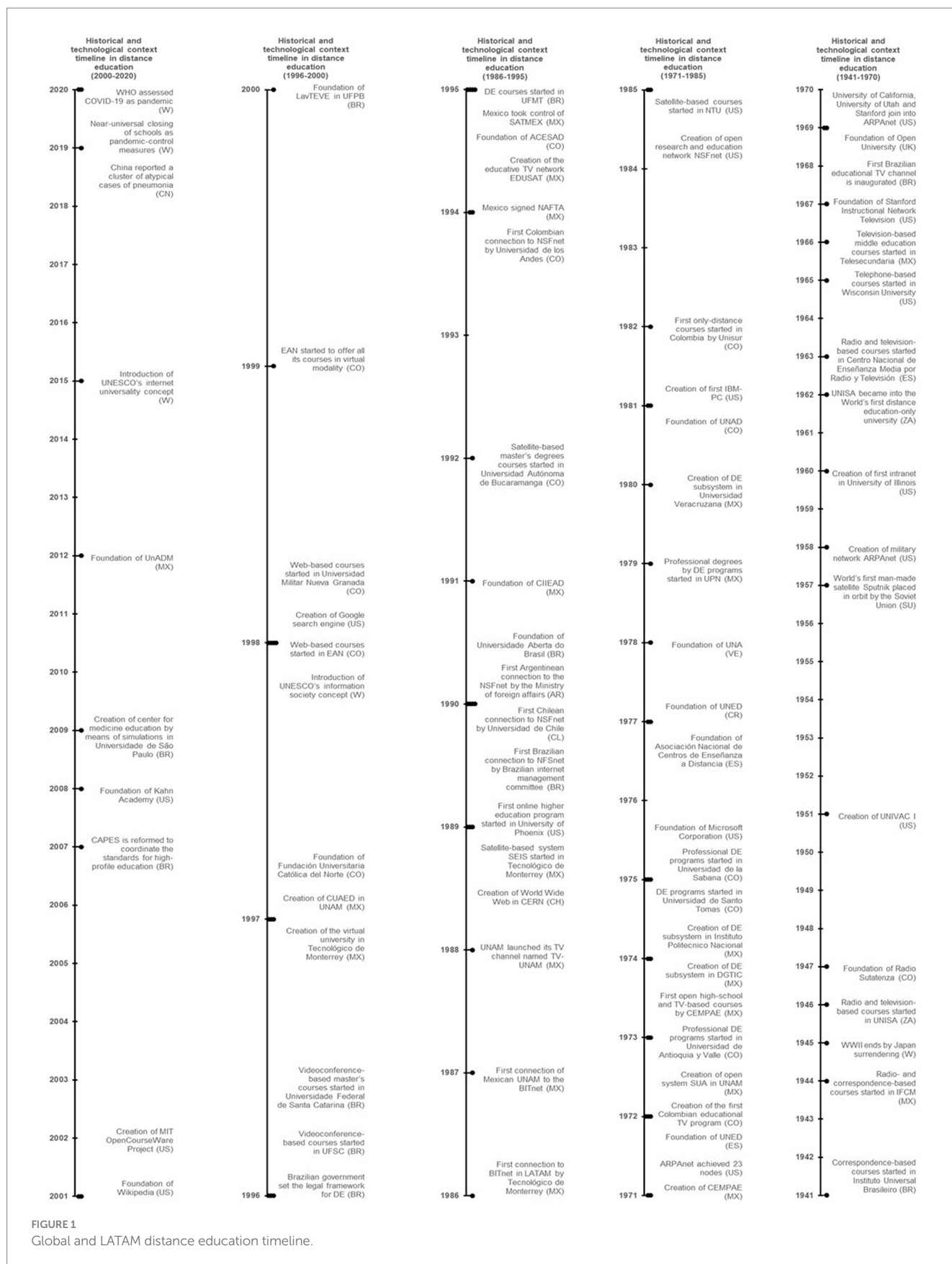


FIGURE 1
Global and LATAM distance education timeline.

the form of personal computers and the Internet. Historically, computers evolved from specific purpose military machines, as the British Turing-Welchman Bomb electro-mechanical

machine designed during the 1940's, to the general purpose commercial digital machines, such as UNIVAC I in 1951 (Tatnall, 2012).

In 1957, the launching of the World's first man-made satellite, Sputnik I, yielded developments in technology, global communications, and education; it was launched by the Soviet Union and triggered an arms race against the US. To fulfil the urgent necessity of high skilled scientists and engineers, positions recurrently covered by foreigners' refugees displaced after World War II, the US implemented a major scientific-technical oriented educational reform (Cha, 2015). Furthermore, due to the arms potential of artificial satellites and other orbital devices, the Global Network was born as a North American military project called ARPAnet. This was created by the Advanced Research Projects Agency (ARPA), in 1958, as a preventive measure against a possible orbital Soviet attack to the US wired and telephone network (Naughton, 2016).

Regarding educational institutions, in 1960, the University of Illinois created an internal network (intranet) to deliver recorded lectures and instructional materials to the students. This early form of computer-aided learning enhanced access to educational resources, facilitating a more flexible learning environment. By 1969, University of California, Stanford and University of Utah had adopted the ARPAnet to accelerate research processes and foster inter-institutional partnerships. With only 23 nodes in 1971, ARPAnet evolved up to the current internet, estimating 30 billion nodes by 2020 (Koenigsberger, 2014) and becoming a critical tool for academic collaboration, communication, and data sharing among HEI.

Countless private initiatives have contributed to shape the internet and other ICT; e.g. the creation of the first graphical user interfaces (GUI) by Xerox PARC, the foundation of the Microsoft Corporation in 1975 by Bill Gates and Paul Allen, or the introduction of the Personal Computer (PC) by International Business Machines (IBM) in 1981 (Koenigsberger, 2014). However, HEI increased their presence on the emerging Global Network establishing by 1985 the first open research and education network, called National Science Foundation Network (NSFnet) (NSFNET, 1990). By participating in global networks, HEI could access a wider array of resources, collaborate more effectively across geographic boundaries, and increase dissemination of knowledge. Besides, the World Wide Web was born, in 1989, with the World's first-ever website on the internet: <http://info.cern.ch/hypertext/WWW/TheProject.html>. This webpage was created at the *Conseil Européen pour la Recherche Nucléaire* (CERN) (Koenigsberger, 2014). These emerging ICT evidenced, for the first time, the relevance of faculty ICT skills to ensure effective integration of technology in knowledge generation as well as in teaching and learning processes.

3.3 RQ3: at what point and through what processes did Information and Communication Technologies (ICT) integrate with distance education (DE)?

ICT were integrated as support to the formative processes, giving the way to the third generation of DE, named ICT-based Education (Yong et al., 2017). It emerged after 1998, when the United Nations Educational Scientific and Cultural Organization (UNESCO) glimpsed the future role of ICT in teaching, leading to 'Knowledge Society' or 'Information Society' (Severin, 2013). This also contributed to the beginning of virtual education, "under principles of teacher-student cooperation and collaboration by means of ICT" (Torres,

2017). This generation witnessed the emergence of Google search engine, in 1998, and the foundation of Wikipedia, in 2001 (Koenigsberger, 2014). In addition, personal computers, office software, internet browsers, multimedia players, electronic encyclopedias and digital storage systems were massively implemented (Abdullahi, 2014).

The current century has seen the development of two more iterations of DE. The fourth generation of DE, web-based education, integrates a wide variety of tools from the Web 2.0. In words of Khan, "web-based programs can provide asynchronous communication to students and the instructor. Likewise, e-mail, listservs, newsgroups, conferencing tools, etc., can jointly contribute to the creation of a virtual community on the Web" (Khan, 1997, p. 6). Additionally, blogs, web conferences, learning platforms and wikis were continuously adopted as well (Yong and Bedoya, 2016). The relevant components in this distance modality, asynchronous electronic learning (e-learning), autonomy, coverage, and economy, represent advantages over face-to-face instruction (Moreno and Maluche, 2017). For instance, in 2001, the Massachusetts Institute of Technology (MIT) launched the OpenCourseWare Project, to provide free education (Abelson, 2008). In 2008, Salman Kahn founded the non-profit web-based educative institution named Kahn Academy; all its content is free and distributed throughout the Net under creative commons license (Bauer, 2013).

The fifth generation, Interactive Education, adopts virtual experiences for teaching. Virtual reality approaches "simulations of real and imagined worlds," through fully immersive and non-immersive modes (Freina and Ott, 2015). Augmented reality, consisting of layers of virtual information over real-life imaging, is also being used for learning purposes (Bower et al., 2014). In addition, artificial intelligence is being continuously implemented to create adaptative educational systems within e-learning platforms (Popenici and Kerr, 2017). Other technological elements have been created specifically for communication and educative purposes within interactive education. Learning management systems (LMS) have been developed to centralize and automate the course administration, to display self-guided e-learning services, to distribute contents, to consolidate web-based courses and beyond (Ellis, 2009). Blended learning (b-learning), mixing asynchronous e-learning content and face-to-face sessions, has been continuously incorporated (Garrison and Vaughan, 2008). Web-based seminars (webinars), personalized learning experiences and other tools from the Web 3.0 are expected to detonate novel learning strategies or even a different DE generation. For instance, personalized learning is promised to revolutionize the education systems through providing equitable outcomes for all learners, despite their learning skills or disabilities (Zhang et al., 2020).

Beyond history and ITC development, Society has shaped new DE learning models and tools. For instance, as a response to the high cost of formal education institutions, governments limitations and borders, Massive Online Open Courses (MOOC) emerged with the aim of providing education to many students for free (Rizzardini et al., 2014). Additionally, a milestone that will mark the future development of DE was the introduction, in 2015, of UNESCO's internet universality concept. Under a human rights-based approach, the level of development of internet universality is evaluated through UNESCO's indicators or ROAM principles, an acronym for: Rights, Openness, Accessibility to all and Multistakeholder participation (Weber, 2015).

Not only asynchronous experiences have been reported. Watts (2016) collected dozens of synchronous implementations of DE, being mostly perceived as positive and even more favorable than the asynchronous alternative. For instance, Rockinson-Szapkiw and Wendt (2015) reported that students enrolled in synchronous courses increased the number of inquiries over those in asynchronous experiences. Also, an improvement in the social cohesion in workgroups in high school courses imparted via videoconference was reported by Mayer et al. (2017). Even though, blended synchronous experiences, with a group split into remote as well as face-to-face modalities, provided insights on complex implementations merging different generations of DE (Wang et al., 2017).

3.4 RQ4: how has the implementation of Information and Communication Technologies (ICT) in Latin American Education progressed over time?

In comparison with North America and Europe, LATAM has adopted internet, ICT, and DE at its own rhythm. The 20th century was characterized by the advent of globalization and the adoption of new technologies, both irrupting in every human activity. In words of Avendaño Castro and Guacaneme Pineda (2016), “it is possible to talk about financial, technological, environmental, economic, cultural, and social globalization.” This close and global-scale interconnection has consequences in all levels of our society. As Severin (2013) accurately synthesizes, “for better or for worse, good and bad news come early than before: science findings, new cures and solutions, discoveries and innovations; but economic crisis, infections, new weapons and control mechanisms too.” Education in LATAM is not an exception. ICT emerged as natural tools to face this new reality, to adapt teaching practice to the global context.

The first connection to the Internet in LATAM took place in Mexico in 1986, only one year after the creation of NSFnet. It joined the main campus of *Tecnologico de Monterrey*, the most recognized private HEI in Mexico, and the School of Medicine of the University of Texas at Austin through the Because It's Time network (BITnet). A year later, the largest Mexican university, *Universidad Nacional Autónoma de México* (UNAM), was connected to the BITnet through the link established previously by *Tecnologico de Monterrey* (Koenigsberger, 2014), obtaining access to NSFnet afterwards. Other Mexican universities obtained temporary connections to the NSFnet, through agreements with these pioneers.

Other countries in LATAM accessed the NSFnet by means of their own national HEI and governmental entities. In 1990, Chile established its first connection to the NSFnet through the *Universidad de Chile*; meanwhile, Brazil did it through the *Comite Gestor da Internet no Brasil*. The same year, Argentina established its first connection to the NSFnet by the management of *Ministerio de Relaciones Exteriores, Comercio Internacional y Culto*. However, other nations took longer to connect to the global network. For instance, Colombia joined the NSFnet until 1994, by the *Universidad de los Andes* (Islas, 2011).

The introduction of the internet during the 1990s revolutionized the education in LATAM, until then dominated by face-to-face models, through the continuous adoption of distance and semi-presential programs. In words of Rama (2016), virtual components

opened the gate to offer a new modality totally virtual, encouraging the intensive use of ICT and novel strategies within the learning and teaching process. ICT promoted the institutional evolution toward the distance modality and eased the entry of new higher-education suppliers, commonly international entities (Vincent-Lacrin, 2011).

By '90s, LATAM nations already had remarkable experiences in terms of higher education under the distance model. For example: *Sistema Universidad Abierta* in Mexico, established in 1972 by UNAM; *Universidad Nacional Abierta* in Venezuela, established in 1978; *Universidad Nacional de Educación a Distancia* (UNED) in Costa Rica, established in 1977; *Unidad Universitaria del Sur* in Colombia, founded in 1982 as the first only-distance programs institution; *Universidad Nacional Abierta y a Distancia*, created in 1981 in Colombia; and *Sistema de Educación Interactiva por Satélite* (SEIS), a satellite-based system introduced in 1989 by *Tecnologico de Monterrey* (Zubieta-García and Rama-Vilate, 2015). These experiences proved the efficient use of ICT in education, motivating other universities to generate DE experiences, like *Universidade Federal de Santa Catarina* in Brazil that started offering videoconference-based master's degrees toward 1996 or UNED in Costa Rica that used e-mail as main support to tutored campaigns by 1995 (Torres and Rama, 2010).

3.5 RQ5: what was the status of distance education (DE) in the three most populated Latin American countries prior to the COVID-19 pandemic?

The adoption of DE in LATAM countries was completely irregular (Rama, 2016), and describing such heterogeneity in deep is outside the scope of this work. Therefore, to understand the characteristics of educational institutions before the outbreak of Covid-19 the following sections cover DE development in the three most populated countries in this region. By analyzing the history of DE in Mexico, Brazil, and Colombia, we aim to gain insights into the challenges and opportunities of distance learning in LATAM, as well as the diversity of contexts and needs that require different pedagogical and technological solutions (Statista, 2023). Table 3 synthesizes the evolution of DE in these countries by the pioneering institutions through successive generations.

3.5.1 Mexico

DE in Mexico was born in 1944, with the *Instituto Federal de Capacitación del Magisterio* (IFCM). Its aim was to train basic education teachers through a radio- and correspondence-based platforms (Zubieta-García and Rama-Vilate, 2015). Up to that moment, as Navarrete-Cazales and Manzanilla-Granados (2017) said with information of Torres-Bodet, 76% of the federal teachers and 86% of the rural ones did not have a professional instruction. In 1966, DE evolved to a closed-circuit television project, broadcasted by microwave, named *Telesecundaria*; it delivered middle education live lessons taught by a remote teacher and a face-to-face mentor (Navarrete-Cazales and Manzanilla-Granados, 2017). With this approach, the Mexican government tried to educate rural population. Even nowadays, *Telesecundaria* is a current strategy of DE implemented, with experiences about history and geography recently reported (Gordillo Fuentes, 2018).

TABLE 3 Milestones achieved by Mexican, Brazilian, and Colombian pioneers in distance education.

	Description	Mexican pioneers	Brazilian pioneers	Columbian pioneers
1st. Generation. Correspondence education	Institutions delivered, periodically, lessons and contents in printed media through the postal services. It offered the students flexibility (in time, place and pace) to learn with refined instructional material (Taylor, 1995).	IFCM was created by the Mexican government, in 1944, to train teachers of basic education by radio- and correspondence-based programs.	Correspondence-based courses were offered to the Brazilian population since the late years of 19th century. By 1941, Instituto Universal Brasileiro offered correspondence-based courses.	The North American institution Hemphill Schools offered training for radio technicians through printed media and kits with practice materials, during 1930s.
2nd. Generation. Audio-visual education	Institutions used several communication technologies, such as television, radio, telephone, video, and telegraph. Main contents were broadcasted, and students were tutored through regular sessions. It allowed the students flexibility to learn with refined and audio-visual instructional material (Taylor, 1995).	In 1966, Telesecundaria offered a hybrid education model through TV. IFCM radio-broadcasted the courses that were previously provided by printed media. CEMPAE was founded, in 1971, to encourage and manage extracurricular programs. In 1973, CEMPAE and Tecnológico de Monterrey, created a TV-based open high-school. During the 1970s, UNAM, IPN, Colegio de Bachilleres and others created their own open systems. By 1979, UPN had offered DE undergraduate programs. UNAM launched TV-UNAM in 1988.	Radio Sociedade do Rio de Janeiro broadcasted educational contents since 1922. In 1937, the Brazilian government launched Radio Mec, an educational radio station. During 1940s, the Capanema reform flexibilized the Brazilian educational system. Commercial television channels broadcasted programs to take Madureza tests during the 1960s. The first educational TV channel was launched in 1967.	Since 1947, Radio Sutatenza broadcasted radio contents for basic education. Open University of Universidad Javeriana offered TV-based courses, since 1972, to train teachers. In 1973, Universidad de Antioquia y Valle already provided professional DE programs.
3rd Generation. ICTs-based education	HEI implemented virtual programs through ICT Web 1.0 or “read-only web”; such as instant messaging, web pages, office software, virtual books, and encyclopedias. It allowed the students flexibility to learn with refined instructional material. Interaction was still very limited.	In 1997, Tecnológico de Monterrey created its virtual university and UNAM, formalized DE just with the creation of CUAED. UDG, IPN, BUAP, UV, UAEH and TecNM created their virtual and distance programs.	Universidade Aberta do Brasil was founded in 1990. In 1996, the Brazilian government set the legal framework for DE. The same year, Universidade Federal de Santa Catarina gave the first videoconference-based course.	Universidad Autónoma de Bucaramanga gave master's degrees, since 1992, by satellite-based courses. In 1998, Universidad Militar Nueva Granada started to offer undergraduate ICT-based programs. Universitaria Católica del Norte, a totally virtual HEI, was founded in 1997.
4th Generation. Web-based education	Institutions incorporated platforms and tools from Web 2.0 or “read-write web”; for instance, e-mail, listservs, newsgroups, conferencing tools, web browsers and web pages. Students had flexibility to learn with highly refined instructional material and they are in continuous communication with instructors.	Completely web-based and virtual universities such as UMEL, UVEG, UNIVIM, Instituto Consorcio Clavijero, UnADM were created	Online courses were included in several HEI, such as Universidade Federal de Pernambuco, Universidade Federal de Minas Gerais, Universidade Federal do Rio Grande do Sul, Universidade Federal de São Paulo, Universidade Anhembi Morumbi, Pontifícia Universidade Católica de Campinas and Centro Universitário Carioca.	In 1998, Escuela de Administración de Negocios implemented the first Colombian web-based course. One year later, all EAN's were converted to virtual modality. Toward 2017, Sistema Nacional de Información de la Educación Superior reported more than 1,191 DE programs for higher levels.
5th Generation. Interactive education	Institutions implement Web 3.0 (“read-write-execute”) tools and artificial intelligence solutions to provide adaptative contents. Also, virtual and augmented realities provide different learning experiences. Besides flexibility, it offers the students tailor-made courses delivered with high interactivity.	Tecnológico de Monterrey reported virtual reality implementations for learning purposes.	Tecnológico de Monterrey reported virtual reality implementations for learning purposes.	Prior to the COVID-19 pandemic, none Colombian institution has been known to implement interactive DE programs.

By 1970s, several institutions had implemented DE programs in Mexico. The *Centro para el Estudio de Medios y Procedimientos Avanzados de la Educación* (CEMPAE) was created in 1971 to encourage, coordinate and control extracurricular programs. Two years later, CEMPAE in collaboration with *Tecnológico de Monterrey* created the first open high-school through TV-based courses. As previously mentioned, in 1972, UNAM, created the open system *Sistema Universidad Abierta*, as an alternative to the traditional modality (Zubieta-García and Rama-Vilate, 2015). Numerous Mexican secondary and HEI started their own open or DE subsystems during the following years; *Instituto Politécnico Nacional* (IPN) and *Dirección General de Institutos Tecnológicos* in 1974, *Colegio de Bachilleres* in 1976, *Universidad Pedagógica Nacional* (UPN) in 1979 and *Universidad Veracruzana* by 1980 (Navarrete-Cazales and Manzanilla-Granados, 2017).

Toward 1987, up to 28 public and private Mexican institutions accounted with their own open or DE programs. In 1988, UNAM launched its TV channel named *TV-UNAM*. Three years later, the Mexican government created *Comisión Interinstitucional e Interdisciplinaria de Educación Abierta y a Distancia*, as a linking mechanism among the national suppliers of open and distance programs (Bosco and Barrón, 2008). Several educational and technological reforms took place during the 1990s, after Mexico signed the North American Free Trade Agreement (NAFTA) and entered globalization. NAFTA motivated the adoption of novel learning models to fulfill the requirements of a more skilled workforce (Torres and Rama, 2010). Concerning communications, in 1995 the Mexican government created *Sistema de Satélites Mexicanos* (SATMEX) to control the satellite system Morelos launched during the 1980s. The same year, *Red de Televisión Educativa* (EDUSAT) was created as an educative TV network.

UNAM formalized DE by 1997 with the creation of *Coordinación de Universidad Abierta y Educación a Distancia* (CUAED) (Navarrete-Cazales and Manzanilla-Granados, 2017). *Tecnológico de Monterrey* has promoted DE with remarkable experiences, like achieving the first connection to the Internet in LATAM and getting its satellite-based system SEIS. In 1997, this institution created its virtual campus, in which, according to its official website, more than 460,000 high-school and undergraduate students have taken online courses.

Nowadays almost every large HEI in Mexico offers distance modalities. Some of them with special relevance are: *Universidad de Guadalajara* (UDG), *Universidad Veracruzana* (UV), *Benemérita Universidad Autónoma de Puebla* (BUAP) and IPN (García Aretio and Ruiz Corbella, 2015). Although only few are completely virtual, this is the case of *Universidad Mexicana en Línea* (UMEL), *Universidad Virtual del Estado de Guanajuato* (UVEG), *Universidad Virtual del Estado de Michoacán* (UNIVIM), *Instituto Consorcio Clavijero* (Rama, 2014). Also, *Universidad Autónoma del Estado de Hidalgo* (UAEH) and *Tecnológico Nacional de México* (TecNM Virtual) offered graduate and posgraduate degrees through virtual models. Recently, in 2012, the Mexican government established the *Universidad Abierta y a Distancia de México* (UnADM) as a completely virtual institution (Sosa Silva and Infante Mendoza, 2017).

The fifth generation or interactive DE is being slowly adopted by some Mexican institutions. For instance, *Tecnológico de Monterrey* has used virtual reality for teaching abstract and complex concepts, such as those in physics and mathematics, in undergraduate programs

(Calderon and Ruiz, 2019; Acosta et al., 2020). Thus, several immersive learning experiences have been reported in recent years by academics of this HEI (Alvarez et al., 2020). Mostly these projects are supported by the NOVUS Fund, an internal initiative to encourage educational innovation.

Currently, about 200,000 students in Mexico take e-learning courses. These represent only 7.0% of the enrolment in higher education and advanced technical studies in Mexico (Aguilar, 2018). The small number of students in the distance model is intimately related to the digital divide and social inequality. As Gómez (2018) reported, based on statistical data from the Mexican government, two of three houses lack economic resources to get internet services and even lack of interest on it. In terms of academic supply, according to official information from *Dirección General de Acreditación, Incorporación y Revalidación* of the Mexican Ministry of Education (SEP), among 22,994 registered and valid degrees in Mexico, only 7.8% of them implement DE, although up to 28.7% apply hybrid models. Most of these programs correspond to graduate and postgraduate levels, with a participation of 60.2 and 27.1%, respectively. Despite the delocalized nature of DE, these programs are highly concentrated in the central region of Mexico, implemented by institutions in Mexico City and neighboring states. By 2019, 37.4% of them were related to management and business (Secretaría de Educación Pública, 2020).

3.5.2 Brazil

Following trends from foreign international institutions, in the late 19th century, Brazil implemented DE through correspondence courses (da Silva et al., 2019). By 1922, Radio Sociedade do Rio de Janeiro broadcasted educative programs on the Brazilian radio (Reifschneider, 2006). During the 1930s, a school radio channel also was visualized as an educational vehicle (Torres, 2019). Formally, Brazilian DE began until 1937 when the government launched the educative radio station named Radio Mec. The next decade, the dictatorship of Getúlio Vargas set the Capanema Reform to ease the promotion through different educational levels. In 1941, Instituto Universal Brasileiro started offering correspondence-based courses. By 1960s, commercial television channels broadcasted programs to take Madureza tests, a set of exams to get elementary and middle education diplomas. In 1967, the first educational TV channel was inaugurated in Recife, followed by several other similar initiatives (Oliveira, 1980), like *TV-Cultura* in São Paulo and *Fundação Maranhense de Televisão Educativa* (Young et al., 2010). Also, during the 1970s, the Brazilian Tele-education Association was created as a non-government institution (Oliveira, 1980). Notable is the case of *Sistema Educativo Radiofônico de Bragança*, in the Brazilian state of Pará, as a broadcast system focused on youth and adult people (Maciel, 2015).

In words of Lupion Torres and Rama (2018), in the 1940–1970 period, “the legal advances were limited to the legalization of the supplementary model of learning.” The adoption of a national policy marked the end of the 1980s decade, but its deployment within the university context was stopped by pressure of the federal administration, with centralizer legislations. In contrast, private institutions and non-government organizations delivered supplementary education courses for elementary and middle levels either through printed-, TV- or satellite- based media (Torres and

Rama, 2010). Even in our days, radio education has been implemented as a DE option. For instance, Patricio reported the experience of 21 schools in the Fortaleza region (Patricio, 2016).

In 1990, the Brazilian government founded the *Universidade Aberta do Brasil* (UAB), establishing guidelines to subsequently implement other DE programs (Aguilar, 2018). By 1994, *Universidade Federal de Mato Grosso* gave the first admission test for a DE program that initiated a year later. In 1996, two events promoted a fast adoption of the Net and the broadening of online education. The first event was the delivering of the internet to the public access, reserved to HEI and governments until then (Aguilar, 2018). The second event was the creation of the Brazilian legal framework for DE, recognizing it as legal, valid, and equivalent to the face-to-face modality (da Silva et al., 2019). 1996 also marked the beginning of videoconference-based courses given by *Universidade Federal de Santa Catarina* (Torres and Rama, 2010).

In the following years, online courses were included in several Brazilian HEI, such as *Universidade Federal de Pernambuco*, *Universidade Federal de Minas Gerais*, *Universidade Federal do Rio Grande do Sul*, *Universidade Federal de São Paulo*, *Universidade Anhembi Morumbi*, *Pontifícia Universidade Católica de Campinas* and *Centro Universitário Carioca* (Torres and Rama, 2010). However, DE did not include elementary education, due to resistance from many educators, politicians, and society (Nobrega and Rozenfeld, 2019). Consequently, the most important implementations of DE remained in higher education. Toward 2002, four HEI supported by interactive media were registered at the Brazilian Ministry of Education (Torres and Rama, 2010).

The first decade of the 21st century witnessed an increment in the supply and enrolment in distance higher education, making Brazil a main actor in the scene of DE in LATAM. For example, UAB covered the whole Brazilian territory with centers dedicated to DE and face-to-face courses. Other Brazilian HEI with remarkable DE modalities are: *Universidade de Brasília*, *Universidade Federal do Rio de Janeiro* (UFRJ) and *Universidade Católica de Brasília*. Moreover, the enrolment in DE institutions increased from 1,682 in 2000, to 369,766 in 2007 (Lupion and Marques, 2013). In the last 15 years, pushed by middle and middle-upper classes, the access of the Brazilian people to higher education has increased substantially. According to Marta Ferreyra et al. (2017), one potential explanation is the expansion in the supply of private institutions over the last few years; from 2007 to 2018, private initiatives offered over 90% of DE programs (Da Silva et al., 2019). In this period, the enrolment in DE programs doubled, from 15.4% of the total students in 2007, to about 33.0% ten years later (da Educação Superior, 2018).

Finally, the current decade glimpses the implementation of the fifth DE generation in Brazil using novel ICT and the introduction of interactive education. In 2007, the Brazilian government reformed the *Coordenação de Aperfeiçoamento de Pessoal de nível superior* (CAPES) to coordinate the standards of the high-profile education. Currently, CAPES promotes DE and innovative experiences with technological resources and ICT. As a result, Brazilian HEI have implemented immersive learning experiences into distance programs. For instance, in 2009, *Faculdade de Medicina da Universidade de São Paulo* inaugurated a center for simulation-based medicine education (Mariani and Pêgo-Fernandes, 2011). Several experiences implementing virtual reality have been reported as well. *Universidade Federal da Paraíba* (UFPB) opened in 2000 the Laboratory of

Technologies for Virtual Teaching and Statistics to develop medical simulators, interactive visualization applications and other interdisciplinary research (Machado and Moraes, 2011).

3.5.3 Colombia

Education in Colombia ventured into distance modalities since the first generation; however, the first Colombian DE experience consolidated until the second generation. During the 1930s, North American institutions offered correspondence training for technical formation; for instance, Hemphill Schools programs to train radio technicians (Alvarez et al., 2000). In 1947, Colombian DE emerged with *Radio Sutatenza*, a radio-broadcasted program for farmers in Sutatenza, Boyacá (Yong et al., 2017); it transmitted until 1994 (Bernal Alarcón, 2012) basic calculus, health, sickness prevention and agricultural production content. Another milestone was achieved in 1972 when *Universidad Javeriana* created a TV program named *Teaching for the New Man*; it trained basic education teachers through the TV network *Inravisión* which belonged to Instituto Nacional de Radio y Televisión (Toro and Rama, 2013). The first Colombian HEI to provide DE degrees was *Universidad de Antioquia y Valle* toward 1973; it focused on the formation of teachers as well, avoiding the daily transportation of the students in the Colombian roads not suitable by then. Two years later, *Universidad de Santo Tomas* and Instituto de Educación a Distancia de la Universidad de la Sabana started their own DE programs for administration, fine arts and physics. Another relevant antecedent of distance higher education in Colombia is the establishment of *Universidad Nacional a Distancia de Colombia* in 1981 (Verdún, 2016).

The government has used DE as a strategy to increase the professional formation of the Colombian people. During the Belisario Betancur's presidency, from 1982 to 1986, DE was officially established as a national policy (Moreno, 2007). Consequently, relevant institutions were created to encourage DE in Colombia; for instance, *Sistema de Educación a Distancia* (SED) (Aguilar, 2018) and *Unidad Universitaria del Sur de Bogotá* (UNISUR), which later became *Universidad Nacional Abierta y a Distancia* (UNAD). Toro and Rama (2013) said "this novel modality departs as a historic opportunity to revamp the national apparatus of education, in its flexibility, organization, financing, administration and pedagogic innovation."

The first implementations of ICT in HEI emerged by 1990s, promoting the Colombian DE evolution toward the third generation. In 1992, *Universidad Autónoma de Bucaramanga* in collaboration with the Mexican institution *Tecnológico de Monterrey* offered satellite-based courses for master's degrees (Aguilar, 2018). In 1995, *Asociación Colombiana de Instituciones de Educación Superior con Programas a Distancia y Virtual* was created. In 1997, *Fundación Universitaria Católica del Norte*, a completely virtual HEI was established, followed by *Universitaria Virtual Internacional* and *Corporación Universitaria de Asturias*, both located in Bogota (Toro and Rama, 2013). Virtual education emerged from all these experiences, and it was defined by the Colombian Ministry of Education as the development of formation programs with the cyberspace as scenery for teaching and learning (Moreno and Maluche, 2017).

Concerning the fourth generation of DE in Colombia, by 1998, the business school *Escuela de Administración de Negocios* (EAN) designed web-based training for teachers using ICT. The same year, the official military institution, *Universidad Militar Nueva Granada*, developed web-based programs. A year later, EAN implemented all its

classes in virtual modality (Yong et al., 2017). Also, was reported the national radio strategy Acción Cultural Popular (ACPO) along the Colombia-Venezuela border. This strategy was based on multimedia, teaching materials, speakers or newspapers (Pérez Pinzón, 2015).

The current century has evidenced the increment in the availability of DE and virtual programs in Colombia, mostly encouraged by the Ministry of Education. Such modality has been favored over face-to-face programs due to reduced cost, intensive use of ICT, larger territorial coverage, student's autonomy, and teacher-student relationship. This last characteristic implies that teachers and professors become guides and advisors through the learning processes. In numbers, the adoption of DE in HEI has increased significantly, passing from 184,136 students in 2010, to 405,218 in 2016 (Tovar et al., n.d.). Another significant effort to incorporate ICT into the Colombian educational system was the governmental initiative *Colombia Aprende* (Severin, 2013). Also, didactic materials for teaching based on audiovisual ICT have been reported (Forero Sandoval and Díaz Buitrago, 2018). Radio, streaming services and social media also have been merged to be used as strategy for international collaboration between HEIs and collaborative radios (Rendón López and Ortega Carrillo, 2015). Toward 2017, according to the information *Sistema Nacional de Información de la Educación Superior* (SNIES), more than 1,191 programs have been registered as distance or virtual ones. Among them, 552 corresponded to undergraduates and 639 to postgraduates (Moreno and Maluche, 2017).

3.6 RQ6: how did the integration of Information and Communication Technologies (ICT) influence the evolution of distance education (DE) in Latin America prior to the COVID-19 pandemic?

Notwithstanding historically the adoption of DE in LATAM countries was completely irregular, technological milestones have propelled DE as mentioned previously for the global context. In Brazil, DE initiatives started in the late 19th century through correspondence courses. Colombia consolidated DE around the 1930s with North American correspondence courses, while Mexico began DE a decade later with radio- and correspondence-based platforms. Technology advances equipped Brazil to advance toward second DE generation as early as the 1920s; Mexico transitioned to television-based DE in the 1960s, while Colombia evolve to radio-based DE in the 1940s and TV-based in the 1970s (see Table 3).

DE has enabled the positive interrelation between larger dissemination of the use of ICTs and the Internet, and better socio-economic conditions. DE in LATAM has usually being delivered at a cost that is substantially lower than that associated with traditional methods (Oliveira et al., 1992); even more, distance higher education in LATAM has steadily increased enrollment since the 1970s (Lupion Torres and Rama, 2018). DE has risen educational level and facilitated training, which in turn improves the quality of human capital, enhances productivity, and contributes to economic development (Balboni et al., 2011; Hanushek and Woessmann, 2012). Balboni et al. (2011) demonstrated that, in synergy, household income and education positively influence Internet usage.

Governments and institutions in LATAM have tried to take advantage of this positive interrelation and have supported DE and connectivity through different public policies. As mentioned previously, the Mexican government tried to educate the rural population through DE; in Brazil, DE began formally with the launch of an educational radio station and, Colombia's government used DE as a strategy to deliver more professional training.

In LATAM, multiple aspects have limited DE dissemination from being considered for decades as a second-class education (Lupion Torres and Rama, 2018) to the appropriation of the Internet, ICT, and electricity (UNESCO Institute for Statistics, 2012). The situation in the three most populated countries in this region has been described; however, other countries show many weaknesses in this aspect (Chañúñez, 2016). Fainholc (2016) identified several problems in the region, from professors without technological skills to areas without connectivity. Similarly, Chávez Maciel (2016) synthesized diverse indicators to trace the adverse scenario in a quantitative manner and to continuously evaluate the adoption and evolution of DE in LATAM.

Digital transformation of education in LATAM was led by supranational organizations, through blended or mixed learning programs in both, the private and the public sector (Lupion Torres and Rama, 2018). Its popularity raised with a smaller costs and dissemination of communication and digital information infrastructure accompanied by an increase in low-income families and professional job market demands. Several Latin-American nations have been incorporating DE under the guidelines of supranational organizations, such as UNESCO (Silvio, 2010) and, even before the COVID-19 outbreak, they proposed new pedagogical models, based on virtual approaches. In general, DE in LATAM evolved since the beginning of the 21st century due to greater innovation, standardization, complexity, assessments options and increased regulation and accreditation (Lupion Torres and Rama, 2018). Argentina, for instance, recently proposed a digital educational system (Rivas, 2018). Moreover, technologies have been tested and adopted in the DE context (Méndez-Ortega et al., 2020). Virtual learning environments, LMS and MOOCs are being designed and adapted to the local context through different initiatives. HEI in Mexico, Brazil, Argentina, Chile, Colombia, El Salvador, Peru and Dominican Republic have participated in MOOCs (Zubieta-García and Rama-Vilate, 2015). Other examples are multinational collaborative workshops, (Amado-Salvatierra et al., 2014), polls-based analyses on the efficiency of learning models, strengths, and weaknesses analyses of e-learning models (Basantes et al., 2018), and new virtual mobility initiatives (Ruiz-Corbella and Álvarez-González, 2014).

Prior to the COVID-19 outbreak, in Mexico nearly every large HEI already offered distance modalities; in Brazil, DE enrolment significantly grew driven by private institutions, while in Colombia, governmental initiatives promoted the incorporation of ICT and DE was widely adopted. To complement an interpretative and explanatory discussion on the impact of ICTs on the development of DE in LATAM, we have introduced other case studies to get further insights into the achievements and challenges faced. For this purpose, we selected Argentina, Peru, Venezuela, and Puerto Rico (see Table 4). We include the latter to expand the sociocultural context to a Caribbean country that is intimately related to the most developed United States of America.

3.6.1 Argentina

The first experiences in DE in Argentina were through correspondence courses. The oldest record data from 1940, considered within the first generation of DE, with the issuance of diplomas for courses for adult training provided by *Escuela Panamericana de Arte* (Cabral, 2022). The second generation was adopted in 1975, with the arrival of radio, television, audiocassettes, and the telephone, which served as mediators in the DE process, making *Universidad de Buenos Aires* the most prominent pioneer (Cabral, 2022). In the 1990s, as responses to policies established by UNESCO and the World Bank, Argentina started to implement ICT into DE and classrooms (OECD, 2015). In the second half of the 1990s, the expansion of coverage and the growth bandwidth of the Argentinian net, along with international policies such as “*Educación para todos*,” led to the inclusion of ICT in DE. *Universidad Nacional de Quilmes* implemented the first virtual classroom (OECD, 2015). Also, *Red Universitaria de Educación a Distancia de Argentina* (RUEDA) was founded to promote and develop the appropriate use of DE resources (Cabral, 2022). By 2015, a significant number of postgraduate and extension programs were recorded as complements to face-to-face lessons. The most representative were the courses recorded by *Universidad Nacional de Quilmes*, targeted to undergraduate programs (OECD, 2015). Similarly, *Universidad Tecnológica Nacional*, through its “*Campus Virtual Global*” platform, provided ICT-based resources to implement distance and blended courses (OECD, 2015). In contrast with the benefic scenario described above, in 2015, only 18% of public institutions had internet access in their DE classrooms, a figure that increased to 33% in 2019 (INDEC, 2019). Up to 2019, according to this research, no fifth generation DE implementation was reported.

3.6.2 Peru

In a dissimilar manner, there are no records of DE in Peru through correspondence courses, i.e., no evidence about the first generation of DE. Additionally, there are no records of educational programs implemented in the DE mode using television, radio, or telephone as tools, as there was no regulatory framework in Peru. The law entitled “*Ley de Educación Superior*,” published in 1983, did not include remote learning as an educational modality (Rama, 2017). The absence of regulatory guidelines discouraged the advancement of DE in Peru. In principle, one of the limitations to implement it was the lack of a regulatory framework that would establish it as a legal and valid modality in the country. Additionally, there was disinterest due to the perception that this type of education was of low quality.

In the 1990s, all LATAM countries began incorporating ICT into education in response to policies established by UNESCO and the World Bank. This move was also driven by the high demand for higher education and competition among universities (OECD, 2015). *Pontificia Universidad Católica del Perú* (PUCP) pioneered online DE projects in 2011, launching its virtual platform, and incorporating postgraduate programs and diplomas (OECD, 2015). Also, *Centro de Investigaciones y Servicios Educativos* and *Facultad de Educación* launched remote programs within PUCP (OECD, 2015). A second DE project was launched by *Sistema de Universidad Abierta* (SUA) in 2014. Thus, various undergraduate programs such as education, law, accounting, business administration, and tourism have been offered in a blended modality (OECD, 2015). By 2015, only two among 75 institutionalized universities in Peru had DE programs, whereas only 12 reported virtual platforms

(OECD, 2015). As in the previous case, no fifth generation of DE experience was reported prior to the COVID-19 pandemic.

3.6.3 Venezuela

Similarly, to the Argentinean case, DE experiences in Venezuela date back to 1960 with correspondence courses offered by the *Instituto de Mejoramiento Profesional del Magisterio y el Instituto Nacional de Cooperación Educativa* (OECD, 2015). There are no records of remote educational programs implemented using television, radio, or telephone as tools. Venezuela started in 1997 with the implementation of ICT as a support resource in face-to-face programs and, in some cases, as a didactic tool in blended education environments. In 2007, the *Proyecto Nacional de Educación Superior a Distancia* was created to establish a regulatory framework that would ensure the quality of DE (OECD, 2015). None fifth generation of DE implementation was reported until 2019.

3.6.4 Puerto Rico

Puerto Rico was chosen as a particular country within LATAM due to its close relationship with the United States of America. As in the previously discussed countries, there are no records of correspondence DE. In 1936, DE made its debut in Puerto Rico with the first radio broadcast of the program “*Escuela al Aire*,” sponsored by the *Departamento de Instrucción Pública* (DIP). This educational program continued until the end of World War II. In 1949, the DIP, through its *Educación Comunal* agency, launched a comprehensive program in this form of DE study, which gained international renown (Robles, 2014). In the 1950s, state radio and television stations emerged, and the DIP started integrating these novel communication media with the educational curriculum. In 1970, “*Abriendo Caminos*” program allowed students to take courses through television and then validate them, obtaining official certificates (Robles, 2014). In the 1990s, the introduction of ICT in Puerto Rico was reflected in the online programs offered by the *Universidad Interamericana*, with over 44 programs up to master’s degrees. Moreover, another HEIs from Puerto Rico, such as *Fundación Ana G. Méndez*, *Universidad de Puerto Rico*, *Universidad del Sagrado Corazón*, *Universidad Politécnica*, *Colegio Universitario Nacional*, among others, encourage online courses as complement to the face-to-face modality (Robles, 2014). Also, *Universidad Internacional Iberoamericana* was introduced to Puerto Rico as a completely online university, i.e., without the option of face-to-face lessons. Similarly, to the other LATAM countries discussed in this subsection, Puerto Rico did not report any fifth generation of DE experience prior to the COVID-19 pandemic (Robles, 2014).

3.7 Limitations

While this work provides valuable insights on remote or distance education (DE) in Latin America (LATAM) for institutions, governments, and policy makers, it is important to acknowledge its limitations.

Firstly, we limited our search to peer-reviewed and indexed literature, primarily cited using Digital Object Identifiers (DOIs). As a result, non-indexed and non-peer-reviewed literature, such as conference proceedings and newspaper reports, were excluded. Additionally, publications without DOIs may have been inadvertently

TABLE 4 Milestones achieved by Argentinian, Peruvian, Venezuelan, and Puerto Rican pioneers in distance education.

	Argentina	Perú	Venezuela	Puerto Rico
1st Generation. Correspondence education	In 1940 was created the adult-targeted distance courses by Escuela Panamericana de Arte (Cabral, 2022).	None registry found.	In 1960, correspondence- based courses, given by Instituto de Mejoramiento Profesional del Magisterio y el Instituto Nacional de Cooperación Educativa, was introduced (OECD, 2015).	None registry found.
2nd. Generation. Audio-visual education	In 1975, Universidad de Buenos Aires started to use radio, TV and cassettes to delivery lessons (Cabral, 2022).	None registry found.	None registry found.	In 1936, the radio program Escuela al Aire from Departamento de Instrucción Pública, through Educación a la Comunidad organism, was launched (Robles, 2014). During the 1950s, Departamento de Instrucción integrated radio, and TV to the study plans. From 1970, it was created the program “Abriendo Caminos” to give TV-based courses (Robles, 2014)
3rd Generation. ICTs-based education	During the 1990s, Universidad Nacional de Quilmes, created the first virtual classroom. It was founded the Red Universitaria de Educación a Distancia de Argentina (Cabral, 2022).	Pontificia Universidad Católica del Perú started the first DE courses (OECD, 2015). In 2014, Sistema de Universidad Abierta started courses for several programs in blended format (OECD, 2015).	In 1997, HEIs started the usage of ICTs in lessons imparted in fase-to-face and blended formats.	ICTs were introduced for up to 44 programs. Fundación Ana G. Méndez, Universidad de Puerto Rico, Universidad del Sagrado Corazón, Pontificia Universidad Católica de Puerto Rico, Universidad Politécnica, and Colegio Nacional Universitario, implemented extension on-line courses (Robles, 2014).
4th Generation. Web-based education	In 2015, posgraduate and extension programs adopted DE. Also, Universidad Nacional de Quilmes offered undergraduate virtual programs (OECD, 2015). Universidad Tecnológica Nacional implemented blended as well as distance courses (OECD, 2015). By 2015, only 18% of public education institutions reported internet connections, increasing this value was increased to 33% to 2019 (INDEC, 2019).	In 2011, HEIs adopted the DE platform known as PUCP, for posgraduate and undergraduate levels (OECD, 2015). By 2015, only 2 HEIs accounted with DE programs, 12 more reported virtual platforms (OECD, 2015).	In 2007 was created the Proyecto Nacional de Educación Superior a Distancia, with the aim to stablish a regulatory framework for DE (OECD, 2015).	Universidad Internacional Iberoamericana was certified, in 2014, as a completely virtual HEI (Robles, 2014).

omitted. Secondly, we conducted the search in English and Spanish only, which may have resulted in the omission of relevant studies published in other languages, particularly from Asia. Thirdly, while this work bases on published literature and offers a comprehensive understanding, firsthand data such as surveys or interviews with

educators, policymakers, or students might have provided deeper and more localized insights. Lastly, while we correlated major historical events and technological advances with the development of DE, detailed discussions about the sociopolitical, economic, and cultural context of every LATAM country were beyond the scope of this paper.

In this sense, a more explanatory and interpretative perspective of the development of DE in LATAM is limited, and there is a risk of drawing overly generalized conclusions, especially when comparing the state of DE in LATAM with that in Europe and the USA.

Beyond this work focus on improvements and innovations of DE in LATAM, we must also recognize DE's multiple associated risks, from mental health issues to data privacy and security concerns. The implementation of ERT amid the COVID-19 lockdown evidenced that the lack of personal interaction and the increasing screen time associated can affect students' mental health and well-being (Hosseini et al., 2023). In turn, this can affect graduates' social and interpersonal skills, mandatory for a successful professional performance (Ehlers, 2020). Most importantly, DE challenges the teaching-learning-evaluation process from multiple perspectives. The absence of physical cues and conducive learning environment does not favor the emotional component for learning, including engagement and motivation (Ferrer et al., 2022). Despite the vast sources of online learning, unreliable sources lead to misinformation and a diminished quality of education. For DE, fairness and efficacy of assessment still represent a challenge (Camacho-Zuñiga et al., 2023), particularly with the broader use of generative artificial intelligence. DE tools do not always fulfill a Universal Instructional Design adding barriers to students with disabilities. Furthermore, DE, delivered through digital platforms, exacerbates risks to the security of personal data and vulnerability against cyber threats (Terry and Rivera Vargas, 2020).

These limitations also highlight potential areas for future research, including the examination of DE evolution in LATAM from alternative perspectives and the exploration of DE in the rest Latin American countries and in other regions of the World.

3.8 Implications

As these lines are written, generative artificial intelligence technologies, such as ChatGPT and other large language models, are rapidly pervading multiple aspects of human activity (Eloundou et al., 2023). Education is not an exception, triggering a discussion on its potential benefits and risks (Allam et al., 2023). Understanding the historical context and current state of DE and ICT in LATAM can provide experiences and study cases for more informed decision-making on the integration and impact of these technologies for a better trained and educated workforce.

These findings pose several implications for practice and policy:

- Governments and institutions in LATAM should prioritize the development and promotion of DE as a viable means to improve social indicators and well-being. This is particularly relevant given the proven efficiency and cost-effectiveness of DE compared to traditional education and the interdependence of educational level with the economic development.
- Policymakers in LATAM must address the digital divide and promote the adoption of ICT, and other general-purpose technologies, in line with ROAM's principles (A framework established by UNESCO and based on Rights, Openness, Accessibility, and Multistakeholder participation). This will enhance the universal right to education and leverage the benefits of remote learning.
- The experience with Emergency Remote Teaching during the pandemic underscores the need to include DE into the emergency policies of every country. This approach has proven effective in ensuring the continuity of education during natural disasters and political and economic crises.

4 Conclusion

Remote or distance education (DE) is a tool for advancing inclusivity, diversity, equity, and accessibility (IDEA) in education, by making learning more accessible, flexible, diverse, and equitable. It leverages technology to remove traditional barriers to education, thus contributing significantly to creating more inclusive educational environments.

This work traced the evolution of DE in Latin America (LATAM) prior to the COVID-19 outbreak, through the case study of some countries for a broad perspective. It contrasted the development of DE in LATAM with global trends, providing a comprehensive overview of the capacities and experiences of educational institutions in the region.

While North America and Europe have led the adoption of DE since the 19th century, LATAM began embracing DE a century later. Higher Education Institutions (HEI) have headed this adoption, driven by public policies and globalization. HEI have also played a crucial role in implementing Information and Communication Technologies (ICT), verifying the synergy between education and technological development observed in the rest of the Globe.

However, the emergence of ICT-based DE programs in LATAM occurred later than in other regions, a delay that can be attributed to the digital divide and social inequality prevalent in these territories. Despite these challenges, DE in LATAM has proven to be an efficient and cost-effective alternative to traditional education increasing the population's educational level. This parameter correlates with economic development; therefore, DE offers a means to combat poverty and improve well-being.

The COVID-19 pandemic, which hit LATAM particularly hard, bolstered emergency remote teaching (ERT) with DE as the primary mode of education delivery. ERT indeed accelerated the evolution of DE worldwide but, at the same time, worsen the disparity between developed and developing regions. Within the Latin American society, ERT also exacerbated other problematics due to the limited connectivity and lack of access to technological devices required to deliver DE in its digital or online modality.

A clear understanding of DE's evolution, status, and future potential is pivotal for the political governance of educational strategies in LATAM. HEIs, governments, and even supranational entities, need a detailed study of cases from several national contexts in paired with an analysis of DE's impact on other societal factors. This knowledge will equip them to establish effective policies and strategies to overcome the long-term educational impact of the COVID-19 pandemic and to integrate artificial intelligence into education in a beneficial manner.

This work provides a historical perspective that is crucial for navigating the current educational challenges in LATAM and planning for a better future in this volatile, uncertain, complex, and ambiguous World toward the Digital Era.

Author contributions

AM and JM-V contributed to conception, design of the study, and organized the database. AM, JM-V, and CC-Z performed the analysis. AM wrote the first draft of the manuscript. JM-V and CC-Z wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Acknowledgments

The authors would like to acknowledge the support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico in the production of this work.

References

- Abdullahi, H. (2014). The role of ICT in teaching science education in schools. *Int. Lett. Soc. Human. Sci.* 8, 217–223. doi: 10.5901/ijser.2013.v3n9p127
- Abelson, H. (2008). The Creation of OpenCourseWare at MIT. *J. Sci. Educ. Technol.* 17, 164–174. doi: 10.1007/s10956-007-9060-8
- Acosta, R. D. A. S., Cooper, E. M. B. H., and Del Valle, J. C. C. (2020). A Methodology to Teach Complex Variable Concepts based on Visualization. Proceedings of the 2020 International Conference on Education Development and Studies, Paris, France. 18–21.
- Aguilar, M. T. (2018). E-learning in Higher Education in Latin America. *Revista historia de la educacion latinoamericana* 20:278. doi: 10.1787/9789264209992-en
- Allam, H., Dempere, J., Akre, V., Parakash, D., Mazher, N., and Ahamed, J., "Artificial Intelligence in Education: an Argument of Chat-GPT Use in Education," 2023 9th International Conference on Information Technology Trends (ITT), Dubai, United Arab Emirates. (2023), pp. 151–156.
- Alvarez, J., Rodriguez, R., and Martinez, M. (2020). Work in progress—Use of immersive videos of virtual reality in education. 2020 IEEE World Conference on Engineering Education (EDUNINE), 15–18 March 2020. 1–4.
- Alvarez, L. J. G., Sfer, A. L., and Plata, L. A. M. (2000). *La educación superior a distancia en Colombia: Visión histórica y lineamientos para su gestión*. Bogotá: Instituto Colombiano para el Fomento de la Educación Superior ICFES.
- Amado-Salvatierra, H. R., Hernandez, R., and Hilera, J. R. (2014). Teaching and promoting web accessibility in virtual learning environments: A staff training experience in Latin-America. 2014 IEEE Frontiers in Education Conference (FIE) Proceedings, 1–4.
- Avendaño Castro, W. R., and Guacaneme Pineda, R. E. (2016). Educación y globalización: Una visión crítica. *Civilizar Ciencias Sociales y Humanas* 16, 191–206. doi: 10.22518/16578953.543
- Balboni, M., Rovira, S., and Vergara, S. (2011) *ICT in Latin America. A microdata analysis*. Santiago, Chile: United Nations Economic Commission for Latin America and the Caribbean (ECLAC).
- Basantes, A. V., Naranjo, M. E., and Ojeda, V. (2018). Metodología PACIE en la Educación Virtual: Una experiencia en la Universidad Técnica del Norte. *Formación Universitaria* 11, 35–44. doi: 10.4067/S0718-50062018000200035
- Bauer, M. (2013). Kahn Academy. *Ref. Rev.* 27, 42–43. doi: 10.1108/09504121311320206
- Bell, R., and Tight, M. (1993). *Open Universities: A British Tradition?* Bristol, UK: ERIC.
- Bernal Alarcón, H. (2012). Radio Sutatenza: Un modelo colombiano de industria cultural y educativa. *Boletín Cultural y Bibliográfico* 46, 4–41.
- Bosco, M. D., and Barrón, H. S. (2008). *La educación a distancia en México: Narrativa de una historia silenciosa. Distance education in Mexico. A narrative of a silent story*. Mexico City: Universidad Nacional Autónoma de México (UNAM).
- Bower, M., Howe, C., McCredie, N., Robinson, A., and Grover, D. (2014). Augmented Reality in education – cases, places and potentials. *Educ. Media Int.* 51, 1–15. doi: 10.1080/09523987.2014.889400
- Cabral, R. G. (2022). *La educación a distancia: características, historia y potencialidad en Argentina*. Available at: <https://repositorio.uca.edu.ar/handle/123456789/13876> (Accessed March 5, 2024).
- Calderon, R. R., and Ruiz, M. A. G. (2019). Learning physics with virtual environment. 2019 IEEE World Conference on Engineering Education (EDUNINE), 1–5.
- Camacho-Zuñiga, C., Peña-Becerril, M., Cuevas-Cancino, M. O., and Avilés-Rabanales, E. G. (2023). Gains from the transition from face-to-face to digital/online modality to improve education in emergency. *Front. Educ.* 8:1197396. doi: 10.3389/feduc.2023.1197396
- Castillo, É. Y., García, N. N., Corredor, C. M., and Malaver, C. E. C. (2017). Evolución de la educación superior a distancia: desafíos y oportunidades para su gestión. *Revista Virtual Universidad Católica del Norte*, 81–105.
- Cha, H. (2015). *Soviet launch of sputnik: Sputnik-inspired educational reform and changes in private returns in America*. Clemson, SC: Clemson University.
- Chan-Núñez, M. E. (2016). La virtualización de la educación superior en América Latina: Entre tendencias y paradigmas. *Revista de Educación a Distancia (RED)* 48, 1–32. doi: 10.6018/red/48/1
- Chanock, K. (2011). A historical literature review of Australian publications in the field of Academic Language and Learning in the 1980s: Themes, schemes, and schisms: Part Two. *J. Acad. Lang. Learn.* 5, A59–A87.
- Chávez Maciel, F. J. (2016). *La inclusión en la educación superior a distancia en América Latina y el Caribe*.
- Crompton, H., Burke, D., Jordan, K., and Wilson, S. (2022). Support provided for K-12 teachers teaching remotely with technology during emergencies: A systematic review. *J. Res. Technol. Educ.* 54, 473–489. doi: 10.1080/15391523.2021.1899877
- da Educação Superior, C. (2018). Notas Estatísticas 2017. *Coordenada por INEP*. Disponível em.
- da Silva, W. B., Amaro, R., and Mattar, J. (2019). Distance Education and the Open University of Brazil: History, Structure, and Challenges. *Int. Rev. Res. Open Distrib. Learn.* 20, 99–115. doi: 10.19173/irrodl.v20i4.4132
- Ehlers, U.-D. (2020) *Future skills: the future of learning and higher education*. Norderstedt: BoD – Books on Demand.
- Ellis, R. K. (2009). Field guide to learning management systems. *ASTD Learn. Circuits*, 1–8. Available at: <http://www.astd.org/NR/> (Accessed March 5, 2024).
- Eloundou, T., Manning, S., Mishkin, P., and Rock, D. (2023). Gpts are gpts: An early look at the labor market impact potential of large language models. *arXiv*. doi: 10.48550/arXiv.2303.10130
- Fainholc, B. (2016). Presente y futuro latinoamericano de la enseñanza y el aprendizaje en entornos virtuales referidos a educación universitaria. *Revista de Educación a Distancia (RED)* 48, 1–22. doi: 10.6018/red/48/2
- Ferrer, J., Ringer, A., Saville, K., Parris, M. A., and Kashi, K. (2022). Students' motivation and engagement in higher education: the importance of attitude to online learning. *High. Educ.* 83, 317–338. doi: 10.1007/s10734-020-00657-5
- Forero Sandoval, J. D., and Díaz Buitrago, E. M. (2018). Implementación de recursos audiovisuales como estrategia de validación pedagógica. *Anagramas Rumbos Sentidos Comun.* 17, 275–291. doi: 10.22395/anrg.v17n33a12
- Freina, L., and Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives. *Int. Sci. Conf. E-Learn. Softw. Educ.* 1, 10–1007.
- García Aretio, L. (1999). Historia de la Educación a Distancia. *Revista Iberoamericana de Educación a Distancia* 2, 8–27.
- García Aretio, L., and Ruiz Corbella, M. (2015). Educación a distancia, la perspectiva de la universidad en una sociedad en red. Una perspectiva desde las regiones de América Latina, El Caribe y América del Norte. *Revista Española de Educación Comparada*:11. doi: 10.5944/rec.26.2015.15819
- Garrison, D. R., and Vaughan, N. D. (2008). *Blended learning in higher education: Framework, principles, and guidelines*. John Wiley & Sons. San Francisco, US.
- Georgiou, I. (2021). The Literature Review as an Exercise in Historical Thinking. *Hum. Resour. Dev. Rev.* 20, 252–273. doi: 10.1177/15344843211004027

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Gómez, L. A. V. (2018). El uso de las TIC para fomentar el autoaprendizaje en la educación en México. *Revista Electrónica de Investigación e Innovación Educativa-REIIE* 1, 6–20.
- Gordillo Fuentes, E. J. (2018). Diseño de estrategias alternativas para el aprendizaje de la historia y geografía. Caso: Telesecundaria en Tabasco/Design of Alternative Strategies for the Knowledge of History and Geography. Case: Telesecundaria in Tabasco. *RIDE Rev. Iberoam. Para Investig. El Desarro. Educ.* 9, 373–387. doi: 10.23913/ride.v9i17.385
- Hanushek, E. A., and Woessmann, L. (2012). Schooling, educational achievement, and the Latin American growth puzzle. *J. Dev. Econ.* 99, 497–512. doi: 10.1016/j.jdeveco.2012.06.004
- Hosseini, S., Camacho, C., Donjuan, K., Pego, L., and Escamilla, J. (2023). Unplugging for Student Success: Examining the Benefits of Disconnecting from Technology during COVID-19 Education for Emergency Planning. *Educ. Sci.* 13:446. doi: 10.3390/educsci13050446
- INDEC (2019). “Acceso y uso de tecnologías de la información y la comunicación. EPH” in *Ciencia y tecnología*, vol. 5 (Argentina: Ministerio de Economía)
- Islas, O. (2011). Los primeros años de Internet en América Latina. *Razón y palabra* 16, 1–16.
- Kanwar, A., and Daniel, J. (2010). “Distance Education and Open Universities” in *International Encyclopedia of Education* (Amsterdam: Elsevier), 404–410.
- Kaye, A. (1981). “Origins and structures” in *Distance teaching for higher and adult education* (London, UK: Routledge), 15–31.
- Kennedy, M. M. (2007). Defining a Literature. *Educ. Res.* 36, 139–147. doi: 10.3102/0013189X07299197
- Khan, B. H. (1997). *Web-based instruction*. Educational Technology. Educational Technology Publications, Englewood Cliffs, 5–18.
- Koenigsberger, G. (2014). *Los inicios de Internet en México*. Universidad Nacional Autónoma de México, Coordinación de la Investigación Científica, Mexico City.
- Lecourt, D. (1988). France: The National Centre for Distance Teaching. *Prospects* 18, 229–238. doi: 10.1007/BF02195520
- Lee, D., Watson, S. L., and Watson, W. R. (2019). Systematic literature review on self-regulated learning in massive open online courses. *Australas. J. Educ. Technol.* 35:3749. doi: 10.14742/ajet.3749
- Lupion, P., and Marques, L. (2013). Educação virtual nas universidades: as contribuições da aprendizagem colaborativa. *Revista História de la Educación Latinoamericana* 19, 175–204. doi: 10.9757/Rhela.19.08
- Lupion Torres, P., and Rama, C. (2018). Distance Education Leaders in Latin America and the Caribbean. *J. Learn. Dev.* 5, 5–12. doi: 10.56059/jl4d.v5i1.286
- Machado, L. S., and Moraes, R. M. (2011). Virtual reality and beyond: The LabTEVE at UFPP in Brazil. *SBC J. Interact. Syst.* 2, 48–51. doi: 10.5753/jis.2011.578
- Maciel, R. A. (2015). Sistema educativo radiofónico de Bragança: educando jovens e adultos na Amazônia Bragantina (1960–1970). *Rev. Tópicos Educ.* 21, 140–165.
- Mariani, A. W., and Pêgo-Fernandes, P. M. (2011). Medical education: Simulation and virtual reality. *São Paulo Med. J.* 129, 369–370. doi: 10.1590/S1516-31802011000600001
- Marta Ferreyra, M., Avitabile, C., Botero Álvarez, J., Haimovich Paz, F., and Urzúa, S. (2017). *At a crossroads: Higher education in Latin America and the Caribbean*. Washington, DC: The World Bank.
- Mayer, G., Lingle, J., and Usselman, M. (2017). Experiences of advanced high school students in synchronous online recitations. *J. Educ. Technol. Soc.* 20, 15–26.
- Méndez-Ortega, M., Malte, A. P., Pozo, A. C., Cabrera, C. S., and Paredes, V. A. (2020). “Higher Education Virtual Classrooms Applied to Learning Processes in Latin America” in *Technology, Sustainability and Educational Innovation (TSIE)*. eds. A. Basantes-Andrade, M. Naranjo-Toro, M. Zambrano Vizuete and M. Botto-Tobar, vol. 1110 (New York: Springer International Publishing), 352–362.
- MEXT. (2020). MEXT: Ministry of Education, Culture, Sports, Science and Technology. Available at: <https://www.mext.go.jp/en/index.htm> (Accessed December 7, 2020)
- Moreno, L. A. S. (2007). Modelo colombiano de educación abierta ya distancia SED. *Revista Historia de la Educación Colombiana* 10, 73–100.
- Moreno, C. F. P., and Maluche, R. B. P. (2017). Caracterización de la educación a distancia y virtual a nivel de pregrado en Colombia 2017. *Revista Ideales* 6, 7–16.
- Naughton, J. (2016). The evolution of the Internet: From military experiment to General Purpose Technology. *J. Cyber Policy* 1, 5–28. doi: 10.1080/23738871.2016.1157619
- Navarrete-Cazales, Z., and Manzanilla-Granados, H. M. (2017). Panorama de la educación a distancia en México. *Revista Latinoamericana de Estudios Educativos (Colombia)* 13, 65–82. doi: 10.17151/rlee.2017.13.14
- Nobrega, F. A., and Rozenfeld, C. C. d. F. (2019). Virtual Reality in the Teaching of FLE in a Brazilian Public School. *Languages* 4:36. doi: 10.3390/languages4020036
- NSFNET. (1990). *En The User's Directory of Computer Networks* (pp. 247–250). Amsterdam: Elsevier.
- OECD. (2015). *E-learning in Higher Education in Latin America*. OECD: París.
- Oliveira, J. B. A. (1980). The Status of Educational Technology in Brazil. *Program. Learn. Educ. Technol.* 17, 210–217. doi: 10.1080/0033039800170403
- Oliveira, B. A., Rumble, J., and Greville, W. (1992). Educación a distancia en América Latina: análisis de costo – efectividad. Available at: <https://api.semanticscholar.org/CorpusID:127492565> (Accessed March 5, 2024).
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:71. doi: 10.1136/bmj.n71
- Patrício, E. (2016). O rádio educativo na percepção de professores da educação básica – O caso rádio escolar do Programa Mais Educação. *Comun. E Soc.* 30, 405–420. doi: 10.17231/comsoc.30(2016).2505
- Pérez Pinzón, L. R. (2015). Tecnología educativa radiofónica en la frontera colombo-venezolana a mediados del siglo XX. *Anu. Colomb. Hist. Soc. Cult.* 42, 145–176. doi: 10.15446/achsc.v42n1.51348
- Petticrew, M., and Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons: Oxford, UK.
- Popenici, S. A. D., and Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Res. Pract. Technol. Enhanc. Learn.* 12:22. doi: 10.1186/s41039-017-0062-8
- Rama, C. A. (2014). University virtualisation in Latin America. *RUSC Univ. Knowl. Soc. J.* 11:32. doi: 10.7238/rusc.v11i3.1729
- Rama, C. (2016). La fase actual de expansión de la educación en línea o virtual en América Latina. *Universidades* 70, 27–39. doi: 10.36888/udual.universidades.2016.70.428
- Rama, C. (2017). La tercera generación de regulaciones de la educación superior a distancia en América Latina. *Revista Diálogo Educativo* 17, 1085–1124. doi: 10.7213/1981-416X.17.054.DS02
- Reifschneider, M. B. (2006). Distance Education in Brazil and in the United States: A Comparative View. *E-Learn. Digital Media* 3, 583–592. doi: 10.2304/elea.2006.3.4.583
- Remenick, L. (2019). Services and support for nontraditional students in higher education: A historical literature review. *J. Adult Contin. Educ.* 25, 113–130. doi: 10.1177/1477971419842880
- Rendón López, L. M., and Ortega Carrillo, J. A. (2015). Aprendizaje colaborativo mediante redes sociales y radio comunitaria web 2.0. *Revista Lasallista de investigación*, 12, 54–65.
- Rivas, A. (2018). Un sistema educativo digital para la Argentina. *Work document* 165, 40–47.
- Rizzardini, R. H., Gütl, C., Chang, V., and Morales, M. (2014). “MOOC in Latin America: Implementation and Lessons Learned” in *The 2nd International Workshop on Learning Technology for Education in Cloud*. eds. L. Uden, Y.-H. Tao, H.-C. Yang and I.-H. Ting (Netherlands: Springer), 147–158.
- Robles, V. (2014). La Educación Presencial, a Distancia y el uso de las TIC en Puerto Rico. *Educación Virtual y a Distancia en Puerto Rico*, 37, 1–12.
- Rockinson-Szapkiw, A., and Wendt, J. (2015). Technologies that assist in online group work: A comparison of synchronous and asynchronous computer mediated communication technologies on students’ learning and community. *J. Educ. Multimedia Hypermedia* 24, 263–279.
- Ruiz-Corbella, M., and Álvarez-González, B. (2014). Virtual Mobility as an Inclusion Strategy in Higher Education: Research on Distance Education Master Degrees in Europe, Latin America and Asia. *Res. Comp. Int. Educ.* 9, 165–180. doi: 10.2304/rcie.2014.9.2.165
- Secretaría de Educación Pública. (2020). DGAIR. Rvoes otorgados por la autoridad educativa federal a las instituciones particulares de educación superior. Available at: <https://www.sep.gob.mx/es/sep1/indicadores> (Accessed October 20, 2020)
- Severin, E. (2013). *Enfoques estratégicos sobre las tics en educación en América Latina y el Caribe*. Oficina Regional de Educación para América Latina y el Caribe, Santiago, Chile.
- Silvio, J. (2010). *Red digital: Revista de Tecnologías de la Información y Comunicación Educativas*. Universidades virtuales en Iberoamérica. Latin America. 1.
- Simonson, M. (2006). Concepciones sobre la educación abierta y a distancia. *Educación abierta y a distancia*. Editorial UOC, 13–48.
- Sosa Silva, S. M., and Infante Mendoza, M. A. (2017). *La internacionalización de la Educación Superior de la modalidad a distancia y el reto de los tiempos virtuales*. El caso de la UnADM.
- Statista. (2023). Online education in Latin America – statistics & facts. Available at: <https://www.statista.com/topics/7239/online-education-in-latin-america/#topicOverview> (Accessed November 20, 2023)
- Tatnall, A. (2012). *History of computers: Hardware and software development*. *Encyclopedia of Life Support Systems*. UNESCO-Eolss Publishers Co Ltd. Isle of Man.
- Taylor, J. C. (1995). Distance education technologies: The fourth generation. *Australas. J. Educ. Technol.* 11:2072. doi: 10.14742/ajet.2072
- Terry, A., and Rivera Vargas, P. (2020). A Critical look at Educational Technology from a Distance Education Perspective. *Digital Educ. Rev.* 37, 208–229. doi: 10.1344/der.2020.37.208-229

- Toro, N. A., and Rama, C. (2013). *La educación superior a distancia y virtual en Colombia: Nuevas realidades*. ACESAD/Virtual Educa. Bogotá.
- Torres, A. N. C. (2017). La educación a distancia como respuesta a las necesidades educativas del siglo XXI. *Revista Academia y Virtualidad* 10:1. doi: 10.18359/ravi.2241
- Torres, I. D. (2019). Curt Lange y Mário De Andrade: Música, radiodifusión pública y política cultural en la década de 1930. *Sociol. Antropol.* 9, 577–595. doi: 10.1590/2238-38752019v9210
- Torres, P., and Rama, C. (2010). *La educación superior a distancia en América Latina y el Caribe*. Santa Catarina: Santa Catarina.
- Tovar, Y. G., Rodgers, N. R., López, D. A. R., and Quintero, W. O. V. (n.d.). *Anuario estadístico de la educación superior colombiana (2010–2016)*, 364.
- UNESCO. (2020). *How is China ensuring learning when classes are disrupted by coronavirus?* UNESCO. Available at: <https://en.unesco.org/news/how-china-ensuring-learning-when-classes-are-disrupted-coronavirus> (Accessed February 19, 2023)
- UNESCO Institute for Statistics. (2012). *ICT in Education in Latin America and the Caribbean. A regional analysis of ICT integration and e-readiness*. Quebec: UNESCO-UIS
- Verdún, N. (2016). Educación virtual y sus configuraciones emergentes: Notas acerca del e-learning, b-learning y m-learning. *Háblame de TIC* 3, 67–88.
- Vincent-Lacrin, S. (2011). La educación superior transfronteriza: Perspectivas y tendencias. *Innovación Educativa* 11
- Wang, Q., Quek, C. L., and Hu, X. (2017). Designing and improving a blended synchronous learning environment: An educational design research. *Int. Rev. Res. Open Distrib. Learn.* 18. doi: 10.19173/irrodl.v18i3.3034
- Watts, L. (2016). Synchronous and asynchronous communication in distance learning: A review of the literature. *Quart. Rev. Dist. Educ.* 17, 23–32.
- Weber, R. H. (2015). *Principles for governing the Internet: a comparative analysis*. UNESCO Publishing. Paris.
- World Bank. (2020). The COVID-19 Pandemic: Shocks to Education and Policy Responses. World Bank. Available at: <https://www.worldbank.org/en/topic/education/publication/the-covid19-pandemic-shocks-to-education-and-policy-responses> (Accessed August 20, 2020)
- Yong, E., and Bedoya, D. (2016). De la educación tradicional a la educación mediada por TIC: Los procesos de enseñanza-aprendizaje en el siglo XXI. Virtual Educa, Puerto Rico. Available at: <http://acceso.virtualeduca.red/documentos/ponencias/puerto-rico/1061-184b.pdf>. (Accessed March 30, 2020)
- Yong, É., Nagles, N., Corredor, C. M., and Malaver, C. E. C. (2017). Evolución de la educación superior a distancia: Desafíos y oportunidades para su gestión. *Revista Virtual Universidad Católica del Norte* 50, 81–105.
- Young, M., Perraton, H., Jenkins, J., and Dodds, T. (2010). *Distance teaching for the third world: the lion and the clockwork mouse*. Routledge: London.
- Zawacki-Richter, O., and Kouroutchkina, A. (2012). The development of distance education in the Russian Federation and the former Soviet Union. *Int. Rev. Res. Open Distrib. Learn.* 13:165. doi: 10.19173/irrodl.v13i3.1165
- Zhang, L., Basham, J. D., and Yang, S. (2020). Understanding the implementation of personalized learning: A research synthesis. *Educ. Res. Rev.* 31:100339. doi: 10.1016/j.edurev.2020.100339
- Zubieta-García, J., and Rama-Vilate, C. (2015). La Educación a Distancia en México: Una nueva realidad universitaria. CDMX: UNAM.

Frontiers in Education

Explores education and its importance for individuals and society

A multidisciplinary journal that explores research-based approaches to education for human development. It focuses on the global challenges and opportunities education faces, ultimately aiming to improve educational outcomes.

Discover the latest Research Topics

[See more →](#)

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact



Frontiers in Education

