2025

EJEL Volume 23, Issue 2

Editor

Heinrich Söbke

Published by Academic Publishing International Limited
Curtis Farm, Kidmore End, Nr Reading, RG4 9AY, United Kingdom
karen.harris@academic-publishing.org

eISSN: 1479-4403

EJEL Volume 23, Issue 2

Prompting Minds: Evaluating how Students Perceive Generative AI's Critical Thinking Dispositions

Luciana Oliveira[1], Célia Tavares[1], Artur Strzelecki[2] and Manuel Silva[1]
[1]CEOS.PP ISCAP Polytechnic of Porto, Portugal
[2]Department of Informatics, University of Economics in Katowice, Poland

Lgo@iscap.ipp.pt (corresponding author)
celiat@iscap.ipp.pt
artur.strzelecki@uekat.pl
mdasilva@iscap.ipp.pt

https://doi.org/10.34190/ejel.23.2.3986

Abstract: As generative artificial intelligence tools like ChatGPT become increasingly integrated into educational environments, understanding their impact on critical thinking is crucial. Despite growing concerns about AI's potential to diminish students' independent reasoning, there is a lack of research tools specifically designed to evaluate students' perceptions of AI's cognitive capabilities. To address this gap, this study introduces the Perceived Critical Thinking Disposition of Generative Artificial Intelligence (PCTD-GAI) scale, aimed at measuring how students perceive generative AI's (GAI) six critical thinking dispositions (reasoning, reaching judgment, search for evidence, search for truth, open-mindedness, and systematicity). While this study validates the scale using ChatGPT, the instrument is adaptable for evaluating other generative AI tools, supporting broader research in AI-driven learning environments, to assess not only how students engage with AI, but also how their reliance on AI may affect their cognitive development and self-regulated learning skills in digital education. To develop and validate the PCTD-GAI scale, the Marmara Critical Thinking Dispositions Scale (MCTDS) was adapted, ensuring relevance to AI assessment while maintaining conceptual robustness. A quantitative cross-sectional study was conducted with 931 university students from Portugal and Poland, employing exploratory and confirmatory factor analyses (EFA & CFA) to assess the scale's validity and reliability. The results demonstrate that the PCTD-GAI effectively captures students' perceptions of ChatGPT's critical thinking dispositions across six key dimensions. Findings indicate moderately positive perceptions across both countries, with Portuguese students consistently rating ChatGPT marginally higher across domains and showing less response variability, suggesting greater consensus. Notably, perceptions were most neutral in the "truth-seeking" domain, while systematicity received the highest ratings, reflecting ChatGPT's perceived systematic capabilities among students. These findings have significant implications for e-learning and AI-driven education. The PCTD-GAI scale enables educators to track students' evolving AI literacy and develop targeted interventions that promote critical AI engagement rather than passive reliance on AI-generated content. Moreover, this research advances the field of e-learning by offering an empirical basis for integrating AI assessment into digital learning strategies, ensuring that AI serves as a cognitive tool rather than a substitute for independent reasoning. The validated PCTD-GAI scale provides a reliable, scalable method for assessing students' perceptions of AI's cognitive capabilities, supporting evidence-based AI pedagogy, and guiding institutional policies on AI integration in education.

Keywords: Critical thinking disposition, Generative artificial intelligence, ChatGPT, Higher education

1. Introduction

The rapid adoption of generative AI (GAI) in education raises concerns about its influence on students' critical thinking. While AI offers innovative learning opportunities, it also poses challenges regarding students' reliance on automated reasoning. This study examines students' perceptions of ChatGPT's critical thinking dispositions, using the proposed Perceived Critical Thinking Disposition of Generative Artificial Intelligence (PCTD-GAI), to better understand how AI tools shape cognitive engagement in academic settings.

Prior research reports mixed findings on AI's impact on students' reasoning. While some studies highlight AI's role in enhancing creativity and cognitive engagement (Hutson and Cotroneo, 2023; Yilmaz and Karaoglan Yilmaz, 2023) others warn of the risks of overreliance, which may weaken independent thought (Zawacki-Richter *et al.*, 2019; Bai, Liu and Su, 2023; Crompton and Burke, 2023; Lo, 2023; Cotton, Cotton and Shipway, 2024). As such, there is no clear understanding of how and why GAI can contribute to or impede critical thinking.

The more generative AI tools are woven into education, the more critical it is for students to develop critical thinking skills and their perceptions of AI's critical capabilities. Students' perceptions of ChatGPT's critical thinking dispositions (e.g., analysing and evaluating information) directly influence their interactions and judgments of what the AI produces (Puig *et al.*, 2019; Ruiz-Rojas, Salvador-Ullauri and Acosta-Vargas, 2024).

Cite this article: Oliveira, L. et al. 2025. "Prompting Minds: Evaluating how Students Perceive Generative AI's Critical Thinking Dispositions", *Electronic Journal of e-Learning*, 23(2), pp 01-18, https://doi.org/10.34190/ejel.23.2.3986

According to recent research by Essel *et al.* (2024) and Zaphir *et al.* (2024), students need to be able to teach AI to think critically and critically evaluate the result quality and reliability of its output. However, these studies are clear also on the point that understanding the impact of AI on critical thinking is still an open question, particularly regarding students' receptiveness to, and dependence on, the cognitive power of AI. Thus, a fundamental question remains unanswered: to what extent do university students perceive GAI, particularly ChatGPT, as possessing critical thinking dispositions, and how does this perception influence their engagement with AI-generated content? By addressing this research question, the study aims to clarify whether students critically assess GAI outputs or develop an overreliance on its perceived reasoning abilities.

In this context, the PCTD-GAI scale fills a critical gap in the literature as no existing instrument specifically measures how students perceive an AI system's cognitive and dispositional traits. Prior studies on AI's impact on education (Essel *et al.*, 2024; Ruiz-Rojas, Salvador-Ullauri and Acosta-Vargas, 2024) have primarily focused on its role as a tool for learning rather than an entity with perceived cognitive dispositions. This study builds upon critical thinking disposition frameworks (Facione *et al.*, 1995; Özgenel and Çetin, 2018) to offer a novel tool that assesses how students judge GAI reasoning, systematicity, and truth-seeking tendencies, an area previously unexplored. The Marmara Critical Thinking Dispositions Scale (MCTDS) was adapted to develop this scale for assessing students' perception about GAI's critical thinking dispositions (Özgenel and Çetin, 2018). The PCTD-GAI differs from prior scales, such as the CCTDI and Yoon's Critical Thinking Disposition (YCTD) instrument, in that it is concerned with students' evaluation of the cognitive and dispositional characteristics of an external entity, rather than the critical thinking tendencies of the individual. This adaptation is crucial because, as Zaphir *et al.* (2024) suggest, students' perceptions of AI's abilities directly influence their reliance on and interaction with these tools.

MCTDS was chosen as the base structure for adaptation for its focus on professional decision-making and adaptability to different contexts. By shifting the focus from human critical thinking dispositions to AI-generated reasoning, this scale provides a new lens to explore the relationship between students and GAI technologies. Moreover, this study contributes to the literature by evaluating the extent to which students trust AI tools like ChatGPT to engage in meaningful critical thinking tasks, a key consideration for educators looking to integrate AI into classrooms

ChatGPT was selected for this study due to its widespread academic adoption, ease of access, and advanced conversational capabilities that distinguish it from other generative AI models (Zawacki-Richter *et al.*, 2019; Zaphir *et al.*, 2024). Unlike domain-specific AI tools such as GitHub Copilot, ChatGPT is designed for general knowledge processing and is extensively used by students across multiple disciplines. Additionally, ChatGPT's open-ended dialogue capabilities allow for more in-depth reasoning and systematic information retrieval, making it an ideal candidate for evaluating AI-generated critical thinking dispositions. While this study focuses on ChatGPT, the PCTD-GPT scale is adaptable to assess students' perceptions of other AI tools, provided they demonstrate comparable reasoning and decision-making abilities.

2. Background

2.1 The Rise of Generative AI

The roots of generative AI (GAI) trace back to the early days of AI research. Alan Turing laid the foundation for AI in the 1950s when he suggested in a research paper the idea of machine intelligence. In fact, the opening sentence of his paper was "Can machines think?"(Turing, 1950, p. 1). In its early days, AI relied on rule-based systems, often referred to as symbolic AI or even logical AI (Smolensky, 1987; Domingos *et al.*, 2016), which operated using predefined rules and logical reasoning.

As computing power advanced, machine learning emerged (Fradkov, 2020), marking a shift from manually programmed rules to statistical methods that allowed AI to learn from data. Supervised and unsupervised learning techniques enabled AI to recognize patterns, make predictions, and improve performance over time. Neural networks, inspired by the structure of the human brain, gained attention but were initially limited by computational constraints (Schmidhuber, 2022).

The real breakthrough came with deep learning (Schmidhuber, 2022), which leveraged large-scale neural networks and powerful hardware to process massive datasets. This era saw the rise of convolutional neural networks (O'shea and Nash, 2015) for image processing and recurrent neural networks (Salehinejad *et al.*, 2017) for handling sequential data. A crucial moment occurred in 2017 with the introduction of the Transformer architecture (Vaswani *et al.*, 2017), which revolutionized natural language processing by enabling models to understand and generate human-like text with unprecedented fluency.

The 2020s have been marked by the rapid proliferation of generative AI, which is a subset of AI where conversational interfaces allow autonomous creation of content in response to natural language prompts (Rashid, Duong-Trung and Pinkwart, 2024). The proliferation of GAI started with OpenAI's GPT series, culminating in GPT-4.5 (Howart, 2025), and continued to show the power of large language models in generating human-like text. GAI also expanded beyond text, producing images, music, and even video through models like DALL·E and MidJourney (Hodges, 2024). As these systems improved, AI-generated content became more accurate and useful.

Today, GAI is used in various fields, and education is one of them. In fact, the rise of GAI is reshaping higher education, particularly in how students approach assignments, research, and learning (Chukwuere, 2024; Riaz and Mushtaq, 2024; Solanke, 2024). Thus, GAI tools, such as ChatGPT from OpenAI, Copilot from Microsoft or even Gemini from Google, have introduced both opportunities and challenges, fundamentally altering academic workflows.

2.2 Critical Thinking

The rapid integration of AI in educational settings, namely GAI, has led to some studies, discussed below, regarding its potential to hinder or enhance students' critical thinking skills. Critical thinking (CT) is often defined as the process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action (Ennis, 2015) or, in a more fundamental conception, "the analytical thinking that underlies all rational discourse and enquiry" (Black, 2012, p. 125). It is a process of careful reasoning and perspective-taking to evaluate statements, ideas, and theories, enabling independent positions based on evidence, crucial for active citizenship and innovation (Vincent-Lancrin, 2024). According to Facione (2013) critical thinkers need to strike a balance between skepticism and open-mindedness in order to avoid falling into the trap of taking information at face value. Thus, encouraging critical thinking in education is fundamental for several reasons, including the development of skills that promote deeper cognitive engagement, facilitate comprehensive analysis, and improve the overall educational experience of students (Todorovska, 2024).

CT has long been considered a vital competency in education and professional contexts (Evens, Verburgh and Elen, 2014; Enciso, Enciso and Daza, 2017; Merfeldaite *et al.*, 2019), often comprising two main components: cognitive skills and dispositions. Although both are critical, it is important to separate critical thinking skills from dispositions for complete assessment and development (Beyer, 1987; Siegel, 2010). Thinking critically means using cognitive processes to analyze, evaluate, and synthesize available information (Merfeldaite *et al.*, 2019). These skills are things like being able to reason, interpret and make decisions (Facione, 1990). However, critical thinking dispositions involve attitudes and habitual ways of behavior that make people use their critical thinking skills habitually. The distinction between skills and dispositions is made in a paper by Facione *et al.* (1995), who define them as a willingness to engage in critical thinking; in other words, dispositions are the internal motivation to apply skills.

Various scholars represent ways wherein skills and dispositions interact in the critical thinking process. Perkins (1985), for example, states that critical thinking cannot be effective without skills and dispositions existing together. Similarly, Ennis (1987); Ennis (1996) stated that being a competent critical thinker requires both the ability to reason and the disposition to use that ability in appropriate contexts. This willingness is an observable quality in that it manifests itself in behaviors, namely in openness to new ideas, curiosity, and persistence in problem-solving (Facione, 2000). Measuring dispositions is therefore essential to understanding an individual's overall capacity for critical thinking, since skills alone do not guarantee consistent application in real world situations (Beyer, 1995; Ennis, 2011; Ennis, 2015).

When it comes to emerging technologies – GAI being a good example – critical thinking dispositions matter all the more (Castaño *et al.*, 2023). Students' perceptions of AI's own intellectual rigor might affect their ability to critically assess and engage with AI outputs as they interact with GAI systems such as ChatGPT, which is one of the most popular ones. According to Gadzella *et al.* (2005) critical thinking dispositions are the most important in making decisions in professional and personal life, which still applies to the manner in which students deal with AI generated information. Whether or not students are likely to critically assess GAI, for example, by cross referencing ChatGPT outputs or questioning its underlying assumptions, depends a great deal on their perceptions of ChatGPT's dispositions, including intellectual curiosity, open mindedness, and systematicity. If students believe ChatGPT has these dispositions at a high level, they may be less likely to use their own critical judgment and miss out on the learning and reflection that could occur. However, if this is not managed appropriately, it may reduce the educational value of GAI integration.

The body of literature on critical thinking, within the field of GAI, is still limited, however, some studies have explored this area. For example, there was a study conducted with computer science students at Bartin University in Turkey that found that students who used ChatGPT in programming had significantly higher computational thinking skills, programming self-efficacy, and motivation compared to students who did not use the tool (Yilmaz and Karaoglan Yilmaz, 2023). Another study conducted with Liberal Art students from Missouri in USA demonstrated the potential of generative AI tools to enhance creativity and innovation in the art and design classroom, helping students to understand the importance of communication and critical analysis skills in the creative process (Hutson and Cotroneo, 2023). Another study explored how ChatGPT affected Ghanaian university students' abilities to think critically, creatively, and reflectively, stating that the study's findings show that incorporating ChatGPT influenced the students' critical, reflective, and creative thinking skills (Essel *et al.*, 2024). The authors even state that “One feasible illustration of the significance of ChatGPT in enhancing critical thinking skills is that it furnishes students with the possibility to engage in dialogues with an AI ChatGPT model that prompts them to think critically” (Essel *et al.*, 2024, p. 9), suggesting that ChatGPT's feedback and guidance, might help them to develop a deeper understanding. Ruiz-Rojas, Salvador-Ullauri and Acosta-Vargas (2024) state that AI integration in higher education has the potential to foster deeper cognitive engagement and facilitate collaboration. However, they also note the importance of teaching students to critically assess AI-generated information. On a less positive note, one study showed that GAI may have an impact on learning and memory capabilities, and also the possible decline of critical thinking skills due to an excessive reliance on AI (Bai, Liu and Su, 2023). Hadi *et al.* (2024) also report on concerns related to the widespread use of AI and its potential to promote superficial learning habits and erode students' social and critical thinking skills. Abbas, Jam and Khan (2024) also reveal that students facing higher academic workload and time pressure, are more likely to use ChatGPT, which leads to the development of tendencies for procrastination, memory loss, and dampening of the students' academic performance.

These studies demonstrate that both positive and negative outcomes are achievable, indicating the absence of universal truths regarding the actual effects of GAI on students. In fact, Essel et al. stated that "the impact of ChatGPT on critical thinking, creative thinking, and reflective thinking in relation to students' learning outcomes remains an area yet to be fully explored and understood” (Essel *et al.*, 2024). Hence, although there is no absolute certainty that GAI can either enhance or diminish the development of students' critical thinking skills, students certainly need critical thinking skills to interact with GAI, both for prompting (input) and for evaluating the quality, accuracy, and relevance of its outputs. Zaphir *et al.* (2024, p. 10) also stated that “Regardless of how intuitive these generative AI become, there will always be a need for students to develop and demonstrate critical thinking skills”.

Moreover, the amount of critical evaluation that GAI's outputs may need can also vary among students, depending on whether students perceived AI as more or less competent in critical tasks and/or their own one's sense of self-efficacy. Thus, students' perception of GAI's critical skills may provide important cues and a prior detection of possible overreliance, which should be avoided. This is to say that a previous evaluation of the student's perception of GAI's critical skills may inform educators of the need to sensitize and train students to develop and employ higher levels of critical assessment of GAI's outputs and increased precision in prompting, before actually introducing GAI in classroom activities. To the best of our knowledge, this issue has not yet been addressed in the literature.

The studies that have been identified as being the most closely related research were three. The first study discusses the philosophical implications of AI’s notion of intelligence, arguing that AI lacks the existential and reflective dimensions of human critical thinking (Leung, 2019). Thus, this study aimed to provide a fresh viewpoint on assessing the ethical difficulties associated to AI as a cultural-philosophical or even a politico-theological model of cognition and its social impact on the activity of "thinking," unlike analysing AI as a technical or technological reality.

The second was a study investigating how critically AI think (Zaphir *et al.*, 2024), done by researchers from the University of Queensland in Australia. The paper examines ChatGPT-4's ability to perform cognitive tasks, such as interpretation, analysis, and explanation. The researchers found that while AI can perform certain cognitive skills (like summarizing or analyzing text), but it struggles with more nuanced cognitive values, such as relevance and depth, without explicit prompt engineering. Tasks that require precision and depth in critical thinking are more challenging for AI, especially when real-world context or personal experience is required. However, AI can perform well on tasks involving straightforward cognitive skills like analysis or explanation. The paper ultimately provides a practical tool for educators to assess and redesign their tasks in a way that fosters critical thinking and is resilient to AI intervention.

The third study focused on the potential of using ChatGPT to assess students' critical thinking in online peer feedback (Tang *et al.*, 2024). This study suggested that ChatGPT demonstrated some ability to assess higher dimensions of critical thinking but showed limitations in assessing the more granular secondary dimensions under the higher dimensions of critical thinking. Nonetheless, neither of these studies incorporated an evaluation of users' perceptions regarding the software's critical thinking capabilities. This gap highlights the need for further investigation into how students conceptualize and perceive the cognitive boundaries of AI technologies, specifically GAI tools like the popular ChatGPT. Analyzing students' perceptions of ChatGPT's critical thinking capabilities might help to address both the opportunities and challenges that come with its use. Moreover, it can create a more informed dialogue, guiding the development of AI technologies in a way that supports students.

2.3 Critical Thinking Disposition

To evaluate students' perceptions of the critical thinking capability of ChatGPT, we propose and validate the Perceived Critical Thinking Disposition of Generative Artificial Intelligence (PCTD-GAI), which is designed to measure students' perceived critical disposition of GAI, specifically using one of the most popular systems, ChatGPT. Over the years, the measurement of critical thinking dispositions has developed to reflect the increasing recognition that if we are to assess not only cognitive skills but also the attitudinal components of critical thinking, then measurement must become more sophisticated. A wide range of scales have been developed to capture these dispositions, each of which contributes to the field in its own way. According to Facione (1990), started in 1990 by Facione, The California Critical Thinking Skills Test (CCTST) was developed as part of the American Philosophical Association's (APA) Delphi Report on critical thinking (Facione, 1990). CCTST is intended to assess the core cognitive skills in critical thinking, such as analysis, inference, evaluation, deductive and inductive reasoning.

The CCTST is a major strength in focusing on critical thinking skills and thus provides a useful assessment of the degree to which people can use cognitive abilities to solve real world problems. Test items call on the test taker to make judgments based on evidence and reasoning, reflecting the complexity of decision-making processes. The CCTST has been used widely in educational settings for which the ability to analyze and evaluate information is essential (Facione, 1991; Facione, Facione and Sanchez, 1994; Frisby and and Traffanstedt, 2003; Tang, 2023). However, CCTST mainly concentrates on skills instead of dispositions. It is effective at gauging cognitive abilities but does not assess the attitudinal components of critical thinking, including how much of a person will be willing to engage his or her critical thinking or how a person will feel about seeking truth and fairness. As a consequence of this gap, the California Critical Thinking Disposition Inventory (CCTDI) was developed to complement the CCTST by assessing dispositions in addition to skills (Facione and Facione, 1994).

Grounded in the Delphi Report, the CCTDI assesses seven key dispositions: truth-seeking, open-mindedness, analyticity, systematicity, critical thinking self-confidence, inquisitiveness, and maturity of judgment (Facione, Facione and Sanchez, 1994). The seven dispositions are evaluated with 75 Likert-scale items that comprise the CCTDI. Truth seeking is wanting to find the most accurate understanding and open mindedness is considering other ways of seeing things. The tendency to approach problems methodically is called systematicity and critical thinking self-confidence is the faith in one's own reasoning abilities (Facione and Facione, 1994). The CCTDI's holistic approach to measuring critical thinking tackles both the cognitive and dispositional edges of the process. The CCTDI is a widely validated and used instrument in critical thinking research and is one of the most influential instruments in such research. Its length has been documented as a possible shortcoming, and there has already been some evidence of cultural limitations when applying it to non-Western populations, demanding changes for other populations (Kökdemir and Dönmez, 2003; Pathak, Dewangan and Mohanty, 2021).

After the success of the CCTDI, Yoon (2014) developed Yoon's Critical Thinking Disposition (YCTD) Instrument, specifically for Korean nursing students. The YCTD is derived from CCTDI to meet the Korean cultural context in healthcare education. The YCTD assesses seven dispositions similar to those of the CCTDI: Prudence, systematicity, intellectual eagerness/curiosity, intellectual fairness, healthy skepticism, objectivity, and critical thinking self-confidence (Shin, Park and Kim, 2015). Cultural relevance was one of the key strengths of YCTD because it addresses the specific needs of Korean nursing students and educational context that would not be met in the other languages (Shin, Park and Kim, 2015). The measure offers a more contextually appropriate framework for measuring dispositions in non-Western cultures where critical thinking may not be expressed in the same way as in the West (Shin, Park and Kim, 2015). The addition of objectivity and prudence as key dispositions in objectivity solidifies the nuanced nature of critical thinking to include professional settings. However, like with the CCTDI, the YCTD is a self-report measure, which may introduce bias based on the respondent's perception of him or herself. Moreover, its emphasis on nursing students can restrict its universality, however it provides a model of culturally adapted critical thinking assessment.

The UF/EMI Critical Thinking Disposition Scale that was developed by researchers at the University of Florida is used to assess critical thinking dispositions of high school students (Kilic and Şen, 2014). In response to the need for a dispositional assessment tool appropriate for younger students, an existing instrument such as the CCTDI is usually used with university students and professionals, this scale was created. The dispositions included in the UF/EMI scale are intellectual curiosity and open-mindedness, which are attitudinal aspects of critical thinking that are important at the time of cognitive development of adolescence (Kilic and Şen, 2014). Targeting high school students provides educators with a tool (UF/EMI scale) to assess tendencies towards critical thinking at an earlier stage, when it is important for long term growth in thinking and in academics. However, as with other self-report measures, the UF/EMI scale is biased by over or under estimation of one's own critical thinking tendencies. Moreover, it is also being specifically engineered for high school students, which means it cannot be used by older and more diverse audiences. Future research is needed to explore its potential adaptions to other educational levels.

In 2018, Özgenel and Çetin developed the Marmara Critical Thinking Dispositions Scale (MCTDS) for the purpose of creating a culturally appropriate instrument for measuring critical thinking dispositions of Turkish educators. (Özgenel and Çetin, 2018) designed this scale to measure teachers' and administrators' critical thinking tendencies, which are influenced by the specific challenges associated with their work environment. Like the CCTDI, the Marmara scale assesses dispositions like open-mindedness, intellectual curiosity, and systematicity (Özgenel and Çetin, 2018). But also stresses adaptability and decision-making, both of which are important for the leadership roles in education. The scale is a useful tool for evaluating how educators utilize critical thinking in their professional circumstances, and how their inclinations to work with complex issues.

Our proposal for the Perceived Critical Thinking Disposition of GAI (PCTD-GAI) is based on the MCTDS. In contrast, the PCTDGAI takes a new angle by centering not on the students' own dispositions but on the students' evaluations of the AI's capabilities. PCTD-GAI was based on MCTDS as it provides the best fit to the particular needs of assessing perceptions of an external entity's (e.g., ChatGPT's) critical thinking dispositions, structural relevance, and cultural adaptability. Unlike previous scales, MCTDS was specifically designed to measure professional dispositions in decision making contexts and target dispositions that are important for educators and administrators (e.g., open-mindedness, flexibility, systematicity, and decision making) that are directly relevant to how students evaluate the quality and reliability of GAI outputs.

The PCTD-GAI does not care about how well students themselves think critically, but rather how they perceive the AI's dispositions: does the AI tackle problems systematically, does it show intellectual curiosity, and can its reasoning be trusted? The MCTDS focus on professional critical thinking dispositions is highly appropriate for the adaptation to this purpose because these are traits that educators and decision makers always evaluate in their own environments.

In addition, the PCTD-GAI required a framework that was oriented towards evaluative dispositions rather than those of intellectual openness, truthiness, and systematicity, which are more closely aligned with how students interact with an AI: making judgements about the AI's reasoning, intellectual openness, truthiness, and systematicity rather than on their own thought processes. In this regard, the MCTDS is easily adaptable for judging how students see the critical thinking dispositions of AI systems such as ChatGPT. This is partly because AI is built with the recognition that critical thinking in professional settings frequently involves assessing external entities and tools, which is essentially the same as evaluating AI. Evaluative dimension of critical thinking dispositions is important when judging not only how well one thinks, but how well another agent (human or AI) thinks.

In addition, since GAI tools are used around the world and in various educational systems, it is critical that the scale used to measure its perceived dispositions is culturally flexible. With a focus on cultural adaptability, MCTDS was developed to provide a more flexible and adaptable framework than previous scales that better informs adaptations to assess how students from different backgrounds perceive AI-generated information.

3. Methodology

3.1 Study Design

The adaptation of the MCTDS to create the PCTD-GAI followed three key stages: item rewording and cognitive shifts, translation, and pilot testing. Similar approaches of critical thinking disposition scale were already reported (Bravo *et al.*, 2020).

1. *Item rewording and cognitive shift***.** First, five types of changes were introduced to reword the original MCTDS items, which assess individuals' own critical thinking dispositions, to assess perceptions of GAI’s critical thinking abilities. The primary adaptation consisted of changing the subject focus of the original MCTDS items, which assessed the individual's own critical thinking abilities, to assess how students perceive GAI’s critical thinking abilities. This shift was essential to match the purpose of the PCTD-GAI scale, which measures students' perceptions of ChatGPT's cognitive dispositions, rather than their own (e.g.: "I analyze the relationships between events, ideas, or problems." -> "I believe that ChatGPT can analyze the relationships between events, ideas, or problems."). A reframing in cognitive action was also needed to reflect GAI’s actions instead of the participant’s own. The adaptation of this assessment shows that the assessment is focused on ChatGPT’s ability to engage in critical thinking behaviors, as perceived by the students (e.g.: "I try to explain problems, situations, or events." -> "I believe that ChatGPT can explain problems, situations, or events.". Contextual adaptations were also necessary in some stances, as the original items needed to fit the context in which GAI operates. To this end, some dispositions relevant to humans, such as emotional or affective responses, were omitted or adjusted in the adapted scale to reflect ChatGPT's capabilities, as these traits are not applicable to AI (e.g.: "I use my mental and affective skills to do or learn something new." -> "I believe that ChatGPT can use mental skills to do or learn something new."). Finally, some of the MCTDS items were subjected to language simplification to ensure that the statements aligned with the specific context of using AI in educational settings (e.g.: "I respect people with different opinions." -> "I believe that ChatGPT can respect opinions from different backgrounds.").
2. *Translation.* After adapting the scale in English, professional translators performed forward and backward translations of the PCTD-GAI into Portuguese and Polish to ensure linguistic and cultural equivalency. The translations were validated by bilingual experts familiar with critical thinking research and AI technologies, ensuring that the scale retained its conceptual meaning across all three languages (English, Portuguese, and Polish).
3. *Pilot testing.* The translated versions of the PCTD-GAI were piloted with a small sample of students (n=20) in both Portugal and Poland to assess the clarity and relevance of the items. Based on feedback from the pilot, minor modifications were made to further refine the language (e.g.: the item "I believe that ChatGPT can ask useful questions to help me." was re-written to be clarified into "I believe that ChatGPT is able to ask questions to better understand a topic or an idea that I present to it."). The experts were provided with the translated versions of the scale and asked to evaluate each item based on its alignment with the constructs being measured, its suitability for the AI context, and its overall appropriateness for cross-cultural application. No further modifications were necessary, as the experts found the items to be clearly worded, contextually relevant, and well-suited for measuring students' perceptions of ChatGPT's critical thinking abilities.

3.2 Participants

The PCTD-GAI survey was approved in May 2024 by the CEOS.PP Ethics Committee (Centre for Organisational and Social Studies of the Polytechnic Institute of Porto) before being distributed to the students. This approval process ensured that all aspects of the survey adhered to ethical standards, including informed consent, confidentiality of participants, and data protection. The Portuguese participants were students from the Polytechnic of Porto: following approval from the relevant ethics committee, an invitation to complete the survey was sent to all six schools of the Polytechnic of Porto, allowing any student from the institution to take part. As a result, 685 students from those six schools completed the survey.

Meanwhile, the Polish participants were students from the University of Economics in Katowice, who were similarly invited to take part. A link to the questionnaire was sent to all students of this university, and ultimately, 246 chose to participate. The study was also approved by the Ethics Committee of the University of Economics in Katowice.

3.3 Instrument

The final PCTD-GAI consists of 27 items measured on a 5-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). Each item corresponds to one of the six main dimensions of critical thinking identified in the Marmara scale: Reasoning, Reaching Judgment, Search for Evidence, Search for Truth, Open-mindedness, and Systematicity (**Table 1 in Appendix 1**). This study followed a quantitative approach, employing a cross-sectional self-administered survey design.

3.4 Data Collection

A total of 931 university students participated in this study (685 from Portugal and 246 from Poland). The sample was selected in a purposive, non-probabilistic manner. Although the proportion of participating students was similar at both institutions, the total number of participants was higher at the Polytechnic of Porto due to its larger student body. Throughout the process, participants were informed of the study's purpose, and informed consent was obtained prior to participation.

3.5 Data Analysis

The prepared research instrument was designed in such a way as to prevent the submission of incomplete responses. Consequently, all recorded answers contained a full set of responses. In the second stage, we examined the variation in the answers and found that in the Portuguese group, 45 people provided responses with a variance of 0.0. This meant they had selected identical responses for every question. Meanwhile, in the Polish group, two responses had a variance of zero, and thus those were removed. As a result, the final sample comprised 640 responses in the Portuguese group and 244 responses in the Polish group.

3.6 Validation

3.6.1 Exploratory factor analysis

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were used to assess the construct validity. After several model iterations, this validation was achieved with a final model consisting of 27 components for the Portuguese sample and 25 items for the Polish sample. IBM SPSS was used to conduct the EFA using the Promax orthogonal rotation method, with Kaiser Normalization to permit the factors to be related (Basto and Pereira, 2012).

A criterion was set to ensure adequate saturation in the factors by eliminating items with factor loadings lower than .30 (Thomas and Hayes, 2021). For the Portugal sample, the Kaiser-Meyer-Olkin (KMO) index was .950, indicating an adequate sample for factor analysis, and Bartlett's test of sphericity was significant (χ^2= 7769.918; df= 351; p < .001), supporting the suitability of the data for the EFA. After removing item A6, the final factor model explained 49.45% of the total variance. Specifically, Dimension 1 accounted for 13.80% of the variance, Dimension 2 for 9.24%, Dimension 3 for 8.61%, Dimension 4 for 8.19%, Dimension 5 for 6.35%, and Dimension 6 for 3.25%. These results suggest a robust and well-defined factor structure in the instrument used.

In the Polish sample, the KMO index was .887, confirming the adequacy of the sample for factor analysis. Bartlett's test of sphericity was significant (χ^2 = 2,221.675; df = 300; p < .001), indicating the data's suitability for EFA. After excluding items R4, A6, and OM1, the final factor model explained 46.73% of the total variance. Dimension 1 explained 10.69% of the variance, Dimension 2 accounted for 9.66%, Dimension 3 for 7.70%, Dimension 4 for 7.27%, Dimension 5 for 7.11%, and Dimension 6 for 4.30% (Table 2). This outcome points to a robust and well-defined factor structure in the instrument.

Table 2: Factor Analysis Results for Portuguese and Polish Samples

Statistical Indices	Portuguese Sample	Polish Sample
KMO	.950	.887
Bartlett's Test of Sphericity (χ^2, df, p)	7769.918, 351, p < .001	2221.675, 300, p < .001
Items Removed	A6	R4, A6, OM1
Total Variance Explained (%)	49.45	46.73
Dimension 1 (%)	13.80	10.69
Dimension 2 (%)	9.24	9.66

Statistical Indices	Portuguese Sample	Polish Sample
Dimension 3 (%)	8.61	7.70
Dimension 4 (%)	8.19	7.27
Dimension 5 (%)	6.35	7.11
Dimension 6 (%)	3.25	4.30
Total Number of Items Retained	27	25

Note: KMO = Kaiser–Meyer–Olkin index for sampling adequacy; Items with factor loadings below .30 were removed (Thomas & Hayes, 2021); Bartlett's Test of Sphericity assesses the suitability of the data for factor analysis. After the items listed under "Items Removed" were excluded, the final model explained the stated percentage of total variance for each sample. Six factors emerged in both samples; their individual contributions (%) to the total variance are shown under "Dimension 1–6.".

3.6.2 Confirmatory factor analysis

The CFA was then performed to assess the discrepancy between the observed and model predicted data, using correlation and covariance matrices. The maximum likelihood method was used for this purpose, based on the assumption that the items are multivariate normally distributed. The validity of this assumption was tested using the Mardia Coefficient, which is considered acceptable if its value is below the result of the formula p(p+2) (Bollen, 1989), where p denotes the number of items in the factor model (27 items in Portuguese version and 25 in Polish version). For Portuguese data, we obtained a value of 10.377 and for Polish data a value of 149.168, which indicates that the matrix is normal.

To evaluate the adequacy of the model for Portuguese data, several indexes have been considered. The first is the CFI coefficient (Comparative Fit Index) was .961, and the NFI coefficient (Normed Fit Index) was .928. These values, ranging from 1 to > .95, suggest a perfect fit to the model (Zhang, Dawson and Kline, 2021). For the RMSEA (Root Mean Square Error of Approximation) index, a value between 0.05 and 0.08 would indicate a reasonable approximation error (Shi *et al.*, 2022). In this study, the result from CFA was .07, within the acceptable range. Finally, the TLI (Tucker-Lewis coefficient) is incremental fit indicator. Value close to 1 indicate a very good fit (Lomax and Schumacker, 2012). In this analysis in CFA, coefficient TLI= .942 was obtained, indicating a very good fit. For Polish data CFI was .952, NFI was .859, RMSEA was .068, and TLI was .930

Regarding the construct reliability and validity, the values for average variance extracted (AVE) should exceed .50 (according to the Fornell-Larcker criterion). In relation to the internal consistency of the measurement scale, Cronbach's alpha and composite reliability were used. The literature mainly point out that this coefficients should be greater than .7, however some authors opt that it could also starts from .6 (Hair *et al.*, 2009). Table 3 presents construct reliability and validity measures for both samples. All indices are at the satisfactory level.

Table 3: Construct reliability and validity

Dimension	Cronbach's alpha*		Composite reliability*		AVE*	
	Portugal	Poland	Portugal	Poland	Portugal	Poland
Reasoning	0.797	0.701	0.800	0.719	0.508	0.501
Reaching judgment	0.765	0.757	0.771	0.772	0.519	0.508
Search for evidence	0.760	0.767	0.766	0.780	0.585	0.584
Search for truth	0.741	0.741	0.741	0.746	0.563	0.559
Open-mindedness	0.748	0.647	0.750	0.667	0.570	0.576
Systematicity	0.689	0.668	0.708	0.704	0.535	0.509

Note. *p-value < .05.

4. Results

To obtain a representative sample for this study, data was collected from participants in Portugal and Poland through online surveys. In Portugal, the survey was publicized across the Polytechnic of Porto and disseminated to all eight schools within the Polytechnic, reaching a wide range of participants from various academic backgrounds. The survey was promoted through institutional channels to encourage broad engagement and a diverse sample. For the Polish sample, the survey was distributed within the University of Economics in Katowice.

The Portuguese sample was collected using LimeSurvey, while the Polish sample was gathered through Google Forms. Different platforms were selected because students from the Polytechnic of Porto are accustomed to LimeSurvey while the students from University of Economics are more familiar with Google Forms. Responses were collected over the month of May 2024.

To evaluate ChatGPT's "Reasoning" abilities, Polish and Portuguese students responded to six statements about assessing relationships, explaining problems, evaluating all aspects of situations, gathering enough information before assessment, challenging presented ideas, and examining the causes of events or problems. For the Polish sample the mean score was 3.1 (SD = 1.1) with a median of 3.0, whereas for the Portuguese sample the mean score was slightly higher, at 3.3 (SD = 1.0) with the same median of 3.0. We observe that both groups rate ChatGPT's reasoning capabilities moderately positively, and that Portuguese students rate it somewhat more positively.

To assess the capabilities of ChatGPT in the area of "Reaching judgment", students from Poland and Portugal answered six statements evaluating the model's ability to categorize information by similarities and differences, to draw new conclusions from given information, to identify and assess risks associated with presented situations, to understand presented problems or ideas, to draw general conclusions from a single idea or event, and to ask appropriate questions to understand the topic or idea presented. The mean score for the Polish and Portuguese samples was 3.6 (median = 4.0). For the Polish sample, the standard deviation was 1.1; for the Portuguese sample, 1.0. The results suggest that students from both countries have a moderately positive view of ChatGPT's capabilities in this domain, with a slightly better consensus from Portuguese students as suggested by a lower standard deviation.

ChatGPT's capabilities in "Search for evidence" were evaluated in Poland and Portugal: students responded to four statements regarding the model's ability to provide credible information and strong evidence when supporting opinions, access information from reliable and diverse sources, seek strong evidence to accept or refute presented ideas or information, and assess the correctness and incorrectness of thoughts and actions presented. In the Polish sample the mean score was 2.9 (SD = 1.1) with median 3.0, and in the Portuguese sample the mean score was slightly higher, 3.1 (SD = 1.1), with median 3.0. The results indicate that both groups perceive ChatGPT as having a neutral to slightly positive ability in this domain, and that Portuguese students slightly favorably evaluate ChatGPT.

Students from Poland and Portugal evaluated ChatGPT's capabilities in the "Search for truth" domain by responding to four statements regarding the model's ability to reflect when evaluating information or ideas, to examine the causes of ideas, events, situations, or problems, to use mental and emotional skills to do or to learn something new, and to cope with problems or events in a realistic way. Polish sample scored a mean of 2.5 (SD = 1.2) and a median of 2.0, the Portuguese sample scored a mean of 2.9 (SD = 1.1) and a median of 3.0. The results show that students from both countries have a neutral to slightly negative perception of ChatGPT in this domain, with Portuguese students having a marginally more positive attitude.

Participants rated four statements about how open-minded they perceived ChatGPT to be in considering others' opinions when solving problems or making decisions, respect individuals with different viewpoints, explain the cause of error or behavior, and viewing situations, ideas, or events from different perspectives. The average rating was 3.3 (median 4.0, SD = 1.2) in the Polish sample. The Portuguese students reported a mean rating of 3.3 with a median of 3.0 and standard deviation of 1.0. In general, students from both countries see ChatGPT as moderately open-minded, but Polish students have slightly higher median ratings and more variable responses than Portuguese ones.

Students rated four statements pertaining to ChatGPT's "Systematicity" to evaluate their perceptions, including the ability to draw conclusions from experienced events or information provided, plan schedules and methods for accomplishing tasks or goals, consider personal values when evaluating presented ideas or events, and infer conclusions about ideas, events, problems or situations based on information supplied. Polish and Portuguese students averaged 3.5 with a median of 4.0. The Polish sample had a standard deviation of 1.2, slightly lower than the standard deviation of 1.0 for the Portuguese sample. The results indicate that, in general, students in both countries perceive ChatGPT as having systematic capabilities, with Portuguese students being slightly less variable in their responses, and therefore more consistent in their perceptions.

5. Discussion

ChatGPT is capable of demonstrating critical thinking-like behaviours, specifically in terms of reasoning, systematicity, and evidence evaluation. However, the way students perceive these capabilities has significant implications for their learning behaviours. Previous work has indicated that students who perceive AI tools as highly competent may be less inclined to perform independent cognitive effort (Bai, Liu and Su, 2023; Cotton, Cotton and Shipway, 2024). Our findings reinforce this concern: students who attribute high critical thinking dispositions to ChatGPT tend to depend more on AI for academic tasks, potentially reducing their engagement in independent critical thinking.

As stated in previous studies, the way students interact with AI is largely based on their belief in the AI's intellectual capabilities and is mediated by their self-efficacy, motivation, and critical thinking awareness (Jia & Tu, 2024). While research suggests that AI could be used to improve creativity and problem-solving skills (Yilmaz and Karaoglan Yilmaz, 2023), it is crucial to ensure that AI is used as a cognitive amplifier rather than a replacement for critical analysis. PCTD-GAI scale is a structured instrument to assess these perceptions, helping educators determine whether students critically evaluate AI-generated content or passively accept it (Ruiz-Rojas, Salvador-Ullauri and Acosta-Vargas, 2024). This is especially important to avoid the formation of superficial learning habits, which might be created by uncritical dependence on AI (Hadi *et al.*, 2024).

However, if students perceive GAI as an autonomous thinker, they may reduce their own cognitive effort and weaker problem-solving abilities (Bonacaro *et al.*, 2024) , which could impact their educational outcomes more broadly (Bai, Liu and Su, 2023). This aligns with studies showing that AI capabilities reshape cognitive learning processes, but their direct influence on critical thinking awareness is limited (Jia and Tu, 2024). As such, educational institutions must design AI-integrated curricula that prioritize active engagement and analytical reasoning rather than passive AI dependence.

One of the most effective ways to mitigate uncritical reliance on AI is through metacognitive training, which teaches students to think about their own thinking processes when engaging with AI tools. Metacognition involves self-monitoring, self-regulation, and awareness of one's cognitive biases (Partalo et al., 2019). If students are trained to critically evaluate AI-generated content, cross-check information, and recognize AI's limitations, they are less likely to rely on it uncritically. The PCTD-GAI scale plays a crucial role in structuring this training by identifying specific cognitive risks associated with AI use. The scale's six dispositions—reasoning, systematicity, search for evidence, search for truth, open-mindedness, and reaching judgment—highlight areas where students may either overestimate or underestimate AI's capabilities.

For instance, students who score ChatGPT highly in "Search for Evidence" may assume its responses are always well-supported, even when they lack proper references. In this case, self-regulation training is needed, where students practice cross-verifying AI claims against peer-reviewed sources, for instance (Uzuntiryaki-Kondakci and Çapa-Aydin, 2013). Similarly, students who rate AI high in "Reasoning" but low in "Search for Truth" may recognize AI's logical structure but fail to detect misleading or biased conclusions. Metacognitive strategies, such as AI bias challenges or bias detection exercises (Ossa, Rivas and Saiz, 2023), can help them develop a more critical stance. If the evaluation produces high scores in "Reasoning" and "Systematicity" students may rely on GAI to solve problems rather than engaging in deep thinking themselves, which hinders the development of higher-order problem-solving abilities. In these cases, guided cognitive scaffolding training may be needed to develop stronger independent reasoning skills (Güss and Wiley, 2007), by, for instance, providing them with AI-generated responses but requiring them to reconstruct arguments, evaluate counterpoints, and refine their reasoning.

By integrating PCTD-GAI scale data with metacognitive training, educators can equip students with the necessary skills to critically assess AI-generated information. This ensures that AI serves as a cognitive enhancer rather than a cognitive replacement, fostering independent thinking in AI-driven learning environments. Thus, the PCTD-GAI scale supports adaptive AI literacy programs by allowing educators to tailor instruction to different student needs. Additionally, longitudinal studies using the scale can track how students' AI literacy evolves, ensuring that educational strategies remain effective over time.

Furthermore, the cross-cultural adaptability of the PCTD GPT, validated with Portuguese and Polish students, demonstrates that the scale can be used in a variety of educational settings. Given that AI integration in education is a global phenomenon, this is particularly important. By focusing on external cognitive evaluation—i.e., the perceived dispositions of AI—this work provides a new perspective on how students and AI technologies interact. Beyond individual classroom settings, the implications of these findings are far-reaching. With AI becoming increasingly integrated into academic curricula, educational institutions are increasingly required not only to teach students how to effectively use AI, but also to ensure they develop critical AI literacy. Specifically, it involves assisting students in appreciating the boundaries of AI, including its inability to address complex or context--dependent cognitive tasks, which have been observed in prior studies (Zaphir *et al.*, 2024). This means that educators must find a balance between harnessing the power of AI and ensuring that students continue to play an active role in the learning process. Without these safeguards, students risk forming a skewed understanding of AI's cognitive autonomy, potentially undermining their ability to think critically in academic and professional contexts.

To address these challenges, future research should explore AI perceptions in diverse educational settings, including those with qualitative methodologies that capture students' reasoning processes in greater depth (Partalo, Skopljak and Mihajlović, 2019). Additionally, longitudinal studies should investigate how prolonged exposure to AI in education influences students' cognitive development over time. Incorporating insights from both quantitative psychometric analysis and qualitative reflections on AI-student interactions can provide a more holistic view of AI's role in higher education.

6. Conclusion

The PCTD-GAI scale fills a significant gap in the literature by offering a validated instrument for assessing students' perceptions of AI's critical thinking dispositions. This study provides novel insights into how students evaluate AI-generated reasoning, systematicity, and truth-seeking abilities, contributing to ongoing discussions about AI's role in higher education (Castaño et al., 2023). As AI integration in education continues to grow, it is imperative that students develop the ability to critically assess AI-generated outputs rather than unquestioningly accept them. The PCTD-GAI scale serves as a foundational tool for educators and researchers to monitor AI reliance in learning environments and develop pedagogical strategies that foster deeper critical engagement with AI technologies. Additionally, the scale offers an empirical foundation for guiding AI literacy programs, ensuring that students are equipped with the metacognitive skills needed to engage critically with AI tools rather than becoming passive consumers of AI-generated content.

Beyond individual classroom applications, this study has broader implications for AI adoption in education policy. As institutions increasingly incorporate generative AI into curricula, there is a growing need to assess how students perceive and interact with these technologies. The PCTD-GAI scale can inform AI literacy initiatives, helping universities and policymakers design evidence-based interventions that encourage responsible AI usage while mitigating cognitive overreliance on AI-driven decision-making.

Despite its contributions, this study has limitations. The validation process was conducted exclusively using ChatGPT, and while the scale is theoretically adaptable to other generative AI models, empirical testing across multiple AI platforms remains necessary. Moreover, the sample was restricted to university students from Portugal and Poland, which may limit the generalizability of the findings to other cultural or educational contexts. Although the cross-cultural adaptability of the PCTD GPT was shown, future studies should increase the number of participants from a wider range of countries and educational systems to verify the universality of the scale. Additionally, the study relied on self-reported perceptions, which may be subject to response bias. Students' evaluations of ChatGPT's critical thinking dispositions could be influenced by factors such as prior experience with AI, personal attitudes towards technology, or even misunderstandings about the tool's capabilities. Finally, the cross-sectional design of the study provides a snapshot of students' perceptions at a single point in time without capturing how these perceptions may evolve with increased exposure to or proficiency with AI tools over time.

Future research could employ mixed-method approaches, including qualitative interviews and experimental studies, to further investigate how students engage with AI in real-world academic settings. Furthermore, a longitudinal design would allow researchers to track how students' perceptions evolve as AI becomes more integrated into education.

AI Statement: The authors declare that they have not used generative or assisted artificial intelligence tools at any stage of the paper's conception and revision. All content presented results exclusively from the author's autonomous work, which guarantees originality, integrity, and compliance with ethical and scientific principles.

Ethics Statement: Prior to data collection, ethical approval was obtained from the Ethics Committee of the Polytechnic Institute of Porto, under the reference number 2024-05-06.

References

Abbas, M., Jam, F.A. and Khan, T.I. (2024) 'Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students', *International Journal of Educational Technology in Higher Education*, 21(1), p. 10. Available at: 10.1186/s41239-024-00444-7

Bai, L., Liu, X. and Su, J. (2023) 'ChatGPT: The cognitive effects on learning and memory', *Brain-X*, 1(3), p. e30. Available at: https://doi.org/10.1002/brx2.30

Basto, M. and Pereira, J.M. (2012) 'An SPSS R-Menu for Ordinal Factor Analysis', *Journal of Statistical Software*, 46(4), pp. 1-29. Available at: 10.18637/jss.v046.i04

Beyer, B.K. (1987) *Practical strategies for the teaching of thinking*. ERIC.

Beyer, B.K. (1995) *Critical Thinking. Fastback 385*. ERIC.

Black, B. (2012) 'An overview of a programme of research to support the assessment of Critical Thinking', *Thinking Skills and Creativity*, 7(2), pp. 122-133. Available at: https://doi.org/10.1016/j.tsc.2012.04.003

Bollen, K.A. (1989) *Structural equations with latent variables (210)*. John Wiley & Sons.

Bonacaro, A., Rubbi, I., Artioli, G., Monaco, F., Sarli, L. and Guasconi, M. (2024) 'AI and Big Data: Current and Future Nursing Practitioners' Views on Future of Healthcare Education Provision', *Studies in health technology and informatics*, 315, pp. 200-204.

Bravo, M.J., Galiana, L., Rodrigo, M.F., Navarro-Pérez, J.J. and Oliver, A. (2020) 'An adaptation of the Critical Thinking Disposition Scale in Spanish youth', *Thinking Skills and Creativity*, 38, p. 100748. Available at: https://doi.org/10.1016/j.tsc.2020.100748

Castaño, J.P., Arnal-Pastor, M., Pagán-Castaño, E. and Guijarro-García, M. (2023) 'Bibliometric analysis of the literature on critical thinking: an increasingly important competence for higher education students', *Economic Research-Ekonomska Istraživanja*, 36(2), p. 2125888. Available at: 10.1080/1331677X.2022.2125888

Chukwuere, J. (2024) 'The future of generative AI chatbots in higher education', *arXiv e-prints*, p. arXiv: 2403.13487.

Cotton, D.R.E., Cotton, P.A. and Shipway, J.R. (2024) 'Chatting and cheating: Ensuring academic integrity in the era of ChatGPT', *Innovations in Education and Teaching International*, 61(2), pp. 228-239. Available at: 10.1080/14703297.2023.2190148

Crompton, H. and Burke, D. (2023) 'Artificial intelligence in higher education: the state of the field', *International Journal of Educational Technology in Higher Education*, 20(1), p. 22. Available at: 10.1186/s41239-023-00392-8

Domingos, P., Lowd, D., Kok, S., Nath, A., Poon, H., Richardson, M. and Singla, P. (2016) 'Unifying Logical and Statistical AI', *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science*. New York, NY, USA. Association for Computing Machinery, pp. 1–11. doi: 10.1145/2933575.2935321.

Enciso, O.L.U., Enciso, D.S.U. and Daza, M.D.P.V. (2017) 'Pensamiento crítico y su importancia en la educación: algunas reflexiones', *Rastros Rostros*, 19(34).

Ennis, R.H. (1987) 'A taxonomy of critical thinking dispositions and abilities', in *Teaching thinking skills: Theory and practice.: Series of books in psychology.* New York, NY, US: W H Freeman/Times Books/ Henry Holt & Co, pp. 9-26.

Ennis, R.H. (1996) 'Critical thinking Upper Saddle River', *NJ: Practice Hall*.

Ennis, R.H. (2011) 'The nature of critical thinking: An outline of critical thinking dispositions and abilities', *University of Illinois*, 2(4), pp. 1-8.

Ennis, R.H. (2015) 'Critical Thinking: A Streamlined Conception', in M. Davies and R. Barnett (eds.) *The Palgrave Handbook of Critical Thinking in Higher Education*. New York: Palgrave Macmillan US, pp. 31-47. doi: 10.1057/9781137378057_2.

Essel, H.B., Vlachopoulos, D., Essuman, A.B. and Amankwa, J.O. (2024) 'ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs)', *Computers and Education: Artificial Intelligence*, 6, p. 100198. Available at: https://doi.org/10.1016/j.caeai.2023.100198

Evens, M., Verburgh, A. and Elen, J. (2014) 'The Development of Critical Thinking in Professional and Academic Bachelor Programmes', *Higher education studies*, 4(2), pp. 42-51.

Facione, N.C. and Facione, P.A. (1994) *The California critical thinking skills test: Test manual*. Millbrae, CA: California Academic Press.

Facione, N.C., Facione, P.A. and Sanchez, C.A. (1994) 'Critical Thinking Disposition as a Measure of Competent Clinical Judgment: The Development of the California Critical Thinking Disposition Inventory', *Journal of Nursing Education*, 33(8), pp. 345-350. Available at: doi:10.3928/0148-4834-19941001-05

Facione, P. (1990) 'A statement of expert consensus for purpose of educational assessment and ınstructions', *Research findings and Recommendations, American Philosophical Association (ERIC Document Reproduction Service No. ED 315 423), Newark, DE*.

Facione, P.A. (1991) 'Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment'.

Facione, P.A. (2000) 'The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill', *Informal logic*, 20(1) Available at: 10.22329/il.v20i1.2254

Facione, P.A. (2013) 'Critical thinking: What it is and why it counts', *Insight assessment*, 1(1), pp. 1-23.

Facione, P.A., xe, nchez, C.A., Facione, N.C. and Gainen, J. (1995) 'The disposition toward critical thinking', *The Journal of General Education*, 44(1), pp. 1-25. Available at: http://www.jstor.org/stable/27797240 (Accessed: 2024/10/18/).

Fradkov, A.L. (2020) 'Early History of Machine Learning', *IFAC-PapersOnLine*, 53(2), pp. 1385-1390. Available at: https://doi.org/10.1016/j.ifacol.2020.12.1888

Frisby, C.L. and and Traffanstedt, B.K. (2003) 'Time and Performance on the California Critical Thinking Skills Test', *Journal of College Reading and Learning*, 34(1), pp. 26-43. Available at: 10.1080/10790195.2003.10850154

Gadzella, B.M., Stacks, J., Stephens, R.C. and Masten, W.G. (2005) 'Watson-Glaser critical thinking appraisal, form-S for education majors', *Journal of instructional psychology*, 32(1), p. 9.

Güss, C.D. and Wiley, B. (2007) 'Metacognition of Problem-Solving Strategies in Brazil, India, and the United States', *Journal of Cognition and Culture*, 7(1-2), pp. 1-25. Available at: https://doi.org/10.1163/156853707X171793

Hadi, M.R., Deng, C., Juho Kim, J., Zhou, P., D. Kwon, Y., Hosny Saleh Metwally, A., Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nacke, L.E. and Hui, P. (2024) 'ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions', *Computers in Human Behavior: Artificial Humans*, 2(1), p. 100027. Available at: https://doi.org/10.1016/j.chbah.2023.100027

Hair, J., Black, W., Babin, B. and Anderson, R. (2009) *Multivariate Data Analysis*. 7 edn. Pearson.

Hodges, T. (2024) 'AI image generation in game design: The Future of AI asset Prompting', *2024 Symposium*. Eastern Washington University. Available at: https://dc.ewu.edu/srcw_2024/ps_2024/p2_2024/37.

Howart, J. (2025) 'When Will ChatGPT-5 Be Released (Feb 2025 Info)'. Available at: https://explodingtopics.com/blog/new-chatgpt-release-date.

Hutson, J. and Cotroneo, P. (2023) 'Generative AI tools in art education: Exploring prompt engineering and iterative processes for enhanced creativity', *2023*, Available at: 10.54517/m.v4i1.2164

Jia, X.-H. and Tu, J.-C. (2024) 'Towards a New Conceptual Model of AI-Enhanced Learning for College Students: The Roles of Artificial Intelligence Capabilities, General Self-Efficacy, Learning Motivation, and Critical Thinking Awareness', *Syst.*, 12, p. 74.

Kilic, H. and Şen, A. (2014) 'Turkish adaptation study of UF/EMI critical thinking disposition instrument', *EGITIM VE BILIM-EDUCATION AND SCIENCE*, 39(176).

Kökdemir, D.Y. and Dönmez, A.T.D. (2003) 'Decision making and problem solving in situations of uncertainty'.

Leung, K.-H. (2019) 'The Picture of Artificial Intelligence and the Secularization of Thought', *Political Theology*, 20(6), pp. 457-471. Available at: 10.1080/1462317X.2019.1605725

Lo, C.K. (2023) 'What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature', *Education Sciences*, 13(4), p. 410. Available at: 10.3390/educsci13040410

Lomax, R.G. and Schumacker, R.E. (2012) *A beginner's guide to structural equation 17 modeling*. Routledge Academic.

Merfeldaite, O., Indrašienė, V., Jegelevičienė, V., Railienė, A., Pivorienė, J., Sadauskas, J., Valavičienė, N. and Penkauskienė, D. (2019) *Society. Integration. Education. International Scientific Conference*.

O'shea, K. and Nash, R. (2015) 'An introduction to convolutional neural networks', *arXiv preprint arXiv:1511.08458*.

Ossa, C.J., Rivas, S.F. and Saiz, C. (2023) 'Relation between metacognitive strategies, motivation to think, and critical thinking skills', *Frontiers in Psychology*, 14.

Özgenel, M. and Çetin, M. (2018) 'Development of the Marmara critical thinking dispositions scale: Validity and reliability analysis', *International Journal of Eurasia Social Sciences*, 9(32), pp. 991-1015.

Partalo, D.N., Skopljak, M.G. and Mihajlović, T.V. (2019) 'Stimulating critical thinking in higher education', *Nastava i vaspitanje*.

Pathak, S., Dewangan, R.L. and Mohanty, S.N. (2021) 'Decision Making Under Uncertainty and Problem Solving', in S. Nandan Mohanty (ed.) *Decision Making And Problem Solving: A Practical Guide For Applied Research*. Cham: Springer International Publishing, pp. 1-12. doi: 10.1007/978-3-030-66869-3_1.

Perkins, D.N. (1985) 'Creativity by design', in A. Costa (ed.) *Developing minds: A resource book for teaching thinking, Vol. 1*. Alexandria, Virginia: Association for Supervision and Curriculum Development, pp. 295-297.

Puig, B., Blanco-Anaya, P., Bargiela, I.M. and Crujeiras-Pérez, B. (2019) 'A systematic review on critical thinking intervention studies in higher education across professional fields', *Studies in Higher Education*, 44(5), pp. 860-869. Available at: 10.1080/03075079.2019.1586333

Rashid, S.F., Duong-Trung, N. and Pinkwart, N. (2024) 'Generative AI in Education: Technical Foundations, Applications, and Challenges', in S. Kadry (ed.) *Artificial Intelligence and Education - Shaping the Future of Learning*. Rijeka: IntechOpen. doi: 10.5772/intechopen.1005402.

Riaz, S. and Mushtaq, A. (2024) *2024 Advances in Science and Engineering Technology International Conferences (ASET)*. 3-5 June 2024. doi: 10.1109/ASET60340.2024.10708721.

Ruiz-Rojas, L.I., Salvador-Ullauri, L. and Acosta-Vargas, P. (2024) 'Collaborative Working and Critical Thinking: Adoption of Generative Artificial Intelligence Tools in Higher Education', *Sustainability*, 16(13) Available at: 10.3390/su16135367

Salehinejad, H., Sankar, S., Barfett, J., Colak, E. and Valaee, S. (2017) 'Recent advances in recurrent neural networks', *arXiv preprint arXiv:1801.01078*.

Schmidhuber, J. (2022) 'Annotated history of modern ai and deep learning', *arXiv preprint arXiv:2212.11279*.

Shi, D., DiStefano, C., Maydeu-Olivares, A. and Lee, T. (2022) 'Evaluating SEM Model Fit with Small Degrees of Freedom', *Multivariate Behavioral Research*, 57(2-3), pp. 179-207. Available at: 10.1080/00273171.2020.1868965

Shin, H., Park, C.G. and Kim, H. (2015) 'Validation of Yoon's Critical Thinking Disposition Instrument', *Asian Nursing Research*, 9(4), pp. 342-348. Available at: https://doi.org/10.1016/j.anr.2015.10.004

Siegel, H. (2010) 'Critical Thinking', in P. Peterson, E. Baker and B. Mcgaw (eds.) *International Encyclopedia of Education*. Oxford: Elsevier, pp. 141-145.

Smolensky, P. (1987) 'Connectionist AI, symbolic AI, and the brain', *Artificial Intelligence Review*, 1(2), pp. 95-109. Available at: 10.1007/BF00130011

Solanke, S. (2024) 'The Prospects of Generative AI in Higher Education', *International Journal of Scientific Research in Engineering and Management*, 8(5).

Tang, L. (2023) *A Comparative Study on Critical Thinking Skills of ISEC and Non-ISEC Teachers in Institutions of Higher Education in the North of China*. University of Montana.

Tang, T., Sha, J., Zhao, Y., Wang, S., Wang, Z. and Shen, S. (2024) 'Unveiling the efficacy of ChatGPT in evaluating critical thinking skills through peer feedback analysis: Leveraging existing classification criteria', *Thinking Skills and Creativity*, 53, p. 101607. Available at: https://doi.org/10.1016/j.tsc.2024.101607

Thomas, C.L. and Hayes, A.R. (2021) 'Using exploratory structural equation modeling to investigate the construct validity of the Critical Thinking Disposition Scale', *Journal of Psychoeducational Assessment*, 39(5), pp. 640-648. Available at: 10.1177/0734282921990564

Todorovska, R. (2024) 'Critical Thinking in the Language Classroom–An Important 21st Century Skill', *International Journal of Education and Philology*, 5(1), pp. 39-48. Available at: https://doi.org//10.69648/BFSI1617

Turing, A.M. (1950) 'I.—Computing Machinery and Intelligence', *Mind*, LIX(236), pp. 433-460. Available at: 10.1093/mind/LIX.236.433

Uzuntiryaki-Kondakci, E. and Çapa-Aydin, Y. (2013) 'Predicting Critical Thinking Skills of University Students through Metacognitive Self-Regulation Skills and Chemistry Self-Efficacy', *Kuram Ve Uygulamada Egitim Bilimleri*, 13, pp. 666-670.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. (2017) 'Attention is all you need', *Advances in neural information processing systems*, 30.

Vincent-Lancrin, S. (2024) 'Critical thinking ', in F. Darbellay (ed.) *Elgar Encyclopedia of Interdisciplinarity and Transdisciplinarity*. Cheltenham, UK: Edward Elgar Publishing, pp. 124-128.

Yilmaz, R. and Karaoglan Yilmaz, F.G. (2023) 'The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation', *Computers and Education: Artificial Intelligence*, 4, p. 100147. Available at: 10.1016/j.caeai.2023.100147

Yoon, J. (2014) 'Development of Critical Thinking Skill Evaluation Scale for Nursing Students', *Journal of Korean Academy of Nursing*, 44(2), pp. 129-138. Available at: 10.4040/jkan.2014.44.2.129

Zaphir, L., Lodge, J.M., Lisec, J., McGrath, D. and Khosravi, H. (2024) 'How critically can an AI think? A framework for evaluating the quality of thinking of generative artificial intelligence', *arXiv Preprint*, Available at: 10.48550/arXiv.2406.14769

Zawacki-Richter, O., Marín, V.I., Bond, M. and Gouverneur, F. (2019) 'Systematic review of research on artificial intelligence applications in higher education – where are the educators?', *International Journal of Educational Technology in Higher Education*, 16(1), p. 39. Available at: 10.1186/s41239-019-0171-0

Zhang, M.F., Dawson, J.F. and Kline, R.B. (2021) 'Evaluating the Use of Covariance-Based Structural Equation Modelling with Reflective Measurement in Organizational and Management Research: A Review and Recommendations for Best Practice', *British Journal of Management*, 32(2), pp. 257-272. Available at: 10.1111/1467-8551.12415

Appendix 1: Table 1

Table 1: PCTD-GPT Scale

Code	MCDTS	PCTD-GPT *Adaptation English*	PCTD-GPT *Translation Portuguese*	PCTD-GPT *Translation Polish*
	Reasoning			
R1	I analyze the relationships between events, ideas or problems.	I believe that ChatGPT can analyze the relationships between events, ideas or problems.	Acredito que o ChatGPT pode analisar as relações entre eventos, ideias ou problemas.	Uważam, że ChatGPT może analizować relacje między wydarzeniami, pomysłami lub problemami.
R2	I try to explain problems, situations or events.	I believe that ChatGPT can explain problems, situations or events.	Acredito que o ChatGPT é capaz de explicar problemas, situações ou eventos.	Uważam, że ChatGPT może wyjaśnić problemy, sytuacje lub zdarzenia.
R3	I evaluate all aspects of a problem, situation or event.	I believe that ChatGPT can evaluate all aspects of a problem, situation or event	Acredito que ChatGPT pode avaliar todos os aspetos de um problema, situação ou evento	Uważam, że ChatGPT może ocenić wszystkie aspekty problemu, sytuacji lub wydarzenia.
R4	I gather enough information before evaluating an idea, problem or situation.	I believe that ChatGPT can gather enough information before evaluating an idea, problem or situation.	Acredito que o ChatGPT consegue reunir informação suficiente antes de avaliar uma ideia, problema ou situação.	Uważam, że ChatGPT może zebrać wystarczającą ilość informacji przed oceną pomysłu, problemu lub sytuacji.

R5	I question an idea, information, problem, event or situation I encounter.	I believe that ChatGPT is able to question an idea, information, problem, event or situation that I present to it.	Acredito que o ChatGPT é capaz de questionar uma ideia, informação, problema, evento ou situação que eu lhe apresente.	Uważam, że ChatGPT może zakwestionować przedstawiony przeze mnie pomysł, informację, problem, wydarzenie lub sytuację.
R6	I investigate the cause of events or problems.	I believe that ChatGPT is able to investigate the cause of events or problems.	Acredito que o ChatGPT é capaz de investigar a causa de eventos ou problemas.	Uważam, że ChatGPT może zbadać przyczynę zdarzeń lub problemów.
	Reaching judgment			
A1	I categorize information about an event, idea or problem according to similarities and differences.	I believe that ChatGPT can categorize information about an event, idea or problem according to similarities and differences.	Acredito que o ChatGPT pode categorizar informação sobre um evento, ideia ou problema de acordo com semelhanças e diferenças.	Uważam, że ChatGPT może kategoryzować informacje o wydarzeniu, pomyśle lub problemie według podobieństw i różnic..
A2	I reach a new conclusion from the general information I have learned.	I believe that ChatGPT can reach a new conclusion from the general information I present to it.	Acredito que o ChatGPT pode chegar a uma nova conclusão a partir da informação geral que eu lhe apresento.	Uważam, że ChatGPT może wyciągnąć nowy wniosek na podstawie ogólnych informacji, które podałem.
A3	I assess the risks I have identified in a situation, problem or event.	I believe that ChatGPT can identify and assess the risks of a situation, problem or event that I present to it.	Acredito que o ChatGPT pode identificar e avaliar os riscos de uma situação, problema ou evento que eu lhe apresento.	Uważam, że ChatGPT może zidentyfikować i ocenić ryzyko związane z sytuacją, problemem lub wydarzeniem, które przedstawiam.
A4	I try to understand a problem, idea or event I encounter.	I believe that ChatGPT can understand a problem, idea or event that I present to it.	Acredito que o ChatGPT pode entender um problema, uma ideia ou um evento que eu lhe apresento.	Uważam, że ChatGPT może zrozumieć problem, pomysł lub wydarzenie, które przedstawiam.
A5	I draw a general conclusion from a single idea, event or situation.	I believe that ChatGPT can formulate a general conclusion from a single idea, event or situation.	Acredito que o ChatGPT consegue formular uma conclusão geral a partir de uma única ideia, evento ou situação.	Uważam, że ChatGPT może wyciągnąć ogólne wnioski z pojedynczego pomysłu, wydarzenia lub sytuacji.
A6	I ask appropriate questions to understand a topic or idea.	I believe that ChatGPT is able to ask questions to better understand a topic or an idea that I present to it.	Acredito que o ChatGPT é capaz de fazer perguntas para melhor compreender um tópico ou uma ideia que eu lhe apresento.	Wierzę, że ChatGPT jest w stanie zadawać pytania w celu lepszego zrozumienia tematu lub pomysłu, który mu przedstawiam.
	Search for evidence			
SE1	I support my opinions with reliable information and strong evidence.	I believe that ChatGPT can support opinions with reliable information and solid evidence.	Acredito que o ChatGPT consegue sustentar opiniões com informações fiáveis e evidências sólidas.	Uważam, że ChatGPT może poprzeć opinie wiarygodnymi informacjami i mocnymi dowodami.
SE2	I obtain information from reliable and diverse sources.	I believe that ChatGPT can obtain information from reliable and diverse sources.	Acredito que o ChatGPT pode obter informações de fontes confiáveis e diversas.	Uważam, że ChatGPT może uzyskać informacje z wiarygodnych i różnorodnych źródeł.
SE3	I look for strong evidence to accept the truth of an idea or information I encounter.	I believe that ChatGPT can look for strong evidence to accept/deny the truth of an idea or information that I present to it.	Acredito que o ChatGPT pode procurar evidências fortes para aceitar/negar a verdade de uma ideia ou informação que eu lhe apresente.	Uważam, że ChatGPT może szukać mocnych dowodów, aby zaakceptować/zaprzeczyć prawdziwości przedstawionego przeze mnie pomysłu lub informacji.

SE4	I evaluate the rightness and wrongness of my thoughts and actions.	I believe that ChatGPT can evaluate what is right or wrong in certain thoughts or actions that I present to it.	Acredito que o ChatGPT pode avaliar o que é certo ou errado em certos pensamentos ou ações que eu lhe apresento.	Uważam, że ChatGPT może ocenić słuszność i błędność myśli i działań, które podaję.
	Search for truth			
ST1	I take my time when evaluating information or ideas.	I believe that ChatGPT is thoughtful when evaluating information or ideas.	Acredito que o ChatGPT é ponderado quando avalia informações ou ideias.	Uważam, że ChatGPT zastanawia się, oceniając informacje lub pomysły.
ST2	I investigate the reasons behind an idea, event, situation or problem.	I believe that ChatGPT can investigate the reasons that support an idea, event, situation or problem that I present to it.	Acredito que o ChatGPT pode investigar as razões que suportam uma ideia, evento, situação ou problema que eu lhe apresento.	Uważam, że ChatGPT może zbadać przyczyny pomysłu, wydarzenia, sytuacji lub problemu, który przedstawiam.
ST3	I use my mental and affective skills to do or learn something new.	I believe that ChatGPT can use mental and emotional skills to do or learn something new.	Acredito que o ChatGPT pode usar competências mentais e emocionais para fazer ou aprender algo novo.	Uważam, że ChatGPT może wykorzystać umiejętności umysłowe i emocjonalne, aby zrobić lub nauczyć się czegoś nowego.
ST4	I deal with problems or events in a realistic way.	I believe that ChatGPT can deal with problems or events in a realistic way.	Acredito que o ChatGPT podem lidar com problemas ou eventos de forma realista.	Uważam, że ChatGPT może poradzić sobie z problemami lub zdarzeniami w realistyczny sposób.
	Open-mindedness			
OM1	I take other people's opinions into account when solving problems or making decisions.	I believe that ChatGPT can take into account the opinions of other stakeholders when solving problems or making decisions.	Acredito que o ChatGPT pode levar em conta a opinião de outros intervenientes ao resolver problemas ou tomar decisões.	Uważam, że ChatGPT może brać pod uwagę opinie innych ludzi przy rozwiązywaniu problemów lub podejmowaniu decyzji.
OM2	I respect people with different opinions.	I believe that ChatGPT can respect opinions from different backgrounds.	Acredito que o ChatGPT consegue respeitar opiniões de diferentes origens.	Uważam, że ChatGPT może szanować opinie z różnych środowisk.
OM3	I explain the reason for a mistake or behavior.	I believe that ChatGPT can explain the reason for a particular mistake or behavior.	Acredito que o ChatGPT pode explicar a razão de um determinado erro ou comportamento.	Uważam, że ChatGPT może wyjaśnić przyczynę błędu lub zachowania.
OM4	I look at situations, ideas or events from different perspectives.	I believe that ChatGPT can look at situations, ideas or events from different perspectives.	Acredito que o ChatGPT pode olhar para situações, ideias ou eventos de diferentes perspetivas.	Uważam, że ChatGPT może spojrzeć na sytuacje, pomysły lub wydarzenia z różnych perspektyw.
	Systematicity			
S1	I draw conclusions from events I have experienced or information I have acquired.	I believe that ChatGPT can draw conclusions from events that have already taken place or from information provided.	Acredito que o ChatGPT pode tirar conclusões a partir de eventos já decorridos ou de informações fornecidas.	Uważam, że ChatGPT może wyciągać wnioski z doświadczonych wydarzeń lub dostarczonych informacji.
S2	I plan when and how I will do something.	I believe that ChatGPT is able to define strategies for accomplishing tasks or goals.	Acredito que o ChatGPT é capaz de definir estratégias para a realização de tarefas ou objetivos.	Uważam, że ChatGPT może zaplanować harmonogram i metodę realizacji zadań lub celów.
S3	I take my own values into account when evaluating ideas or events.	I believe that ChatGPT can consider personal values when evaluating	Acredito que o ChatGPT pode considerar valores pessoais ao avaliar	Uważam, że ChatGPT może wziąć pod uwagę wartości osobiste podczas

		ideas or events presented.	ideias ou eventos apresentados.	oceny przedstawionych pomysłów lub wydarzeń.
S4	I make inferences about an idea, event, problem or situation.	I believe that ChatGPT is able to make inferences about an idea, event, problem or situation based on the information presented.	Acredito que o ChatGPT é capaz de fazer inferências sobre uma ideia, evento, problema ou situação com base nas informações apresentadas.	Uważam, że ChatGPT może wyciągać wnioski na temat pomysłu, zdarzenia, problemu lub sytuacji na podstawie dostarczonych informacji.

Riding The Wave of Technology Integration by Applying Student-Centered, Blended Learning Course Design Principles in a Human Anatomy Course

Helga Brigritta Hambrock and Daniële Kristen
Sefako Makgatho Health Sciences University, South Africa

Helga.Hambrock@smu.ac.za
Daniele.Kristen@smu.ac.za

https://doi.org/10.34190/ejel.23.2.3913

Abstract: This case-study focuses on improving the teaching and learning experience of second year students in a Human Anatomy course at a Health Sciences University in South Africa. The students were underperforming, and the repeaters caused classes to become overcrowded. In search of a solution an instructor of the course and an instructional designer at the University worked on a re-design of the course by including technology and increasing interaction with the course content and peer collaboration as supported by research and best practices. A mixed-methods research methodology was utilized by collecting qualitative and quantitative data. The qualitative data was collected from an interview with the lecturer and a feedback survey from the students. For the analysis of the qualitative data a thematic analysis was applied to identify themes and subthemes which were sorted under the following sub-questions: what were the challenges, successes and suggestions of the newly designed course? Additionally, quantitative data was collected from the students' grades. The application of technology and increased student engagement proved to be successful, and the study proposes a framework based on best practices and feedback from the lecturer and students for an improved course design for the future.

Keywords: Case-study, Student performance, Course design, Mixed method, Thematic analysis, Instructional design, Teaching and learning experience, Collaborative learning, Gamification and student-centered learning

1. Introduction

During a time where waves of new technologies are flooding the market, educators are often overwhelmed and rather reluctant to "ride the wave". It may however be a time when these new technologies could be used to offer opportunities to lecturers and students to change their traditional ways of teaching and learning by incorporating student-centered course designs and pedagogies, - These changes could ultimately improve students' interest and performance (Garcia, 2000; Koehler, et.al.,2004; Baldwin, S., Ching, Y. H., and Hsu, Y. C. 2018).

The concerning performance of students in a second year Human Anatomy course at a University of Health Sciences in South Africa the instructor of the class and the instructional designer at the University explored best practices for redesigning the Human Anatomy course. The design focused on moving from traditional lecturer centered behaviorist teaching approach to a student-centered blended learning approach based on best practices. These were based on the ADDIE Instructional designer model and the first principles of Merril (2021). These best practices were used to review the "old" course, and an interactive learning experience was designed, developed, implemented and evaluated. (Figure 1)

2. The Objective of the Study

The aim of this study was to improve the performance of students by changing a traditional lecturer centered course to a student centered interactive, blended learning experience. After the course was redesigned and implemented, the feedback of lecturers and students will be used to compile a framework for future updated course designs.

3. Problem Statement

The problem statement of the study positions that: "Student performance in an identified undergraduate Human Anatomy course needs improvement". This statement leads to the research question: "How can courses be designed to improve student performance in a Human Anatomy course?" To further the investigation the following sub-research questions were included in the interview: "What were the challenges of the old class?"

Cite this article: Hambrock, H.B. and Kristen, D. 2025. "Riding The Wave of Technology Integration by Applying Student-Centered, Blended Learning Course Design Principles in a Human Anatomy Course", *Electronic Journal of e-Learning*, 23(2), pp 19-29, https://doi.org/10.34190/ejel.23.2.3913

"What in your opinion were the successes of the re-designed class"? and "What suggestions do you have for future course-designs for your class?"

4. Background and Literature Review

According to (Brown and Katz; 2011), one should prioritize design in all fields of life by inviting designers to boardrooms and financial meetings, and to "rather enlist designers, to make an already developed idea more attractive". From the instructional design perspective, one needs to apply best practices to achieve the best results (National Research Council [NRC] 2004).

Various theorists in this field state that the design process is a multileveled, collaborative, iterative and theory driven process. (Brown, 1992; Collins, 1999). They argue that theory needs to be tested in a real-world setting which includes schools or universities, districts and communities. By applying the action research: "design-analysis-re-design" iterative research approach, improvements can be made on all levels of a course.

For this study the instructor of the course and the instructional designer collaborated on designing a student-centered approach by adding activities such as 1. Video recorded assignments, 2. Pre-and post quizzes (Mazur, 2013). 3. Applying a flipped classroom approach (Tucker, 2012) as well as implementing 4. Gamification principles (Groh, 2012). These activities increased the student engagement with the content and were respectively introduced into the courses per term. The design of the lessons was based on Merrill's design of instruction (2012) and on the ADDIE instructional design models.

5. Research Methodology

A mixed methods action research methodology was followed for this case study. An action research methodology was selected as the research took on a cyclical iterative process (Figure 1). It refers to a research approach where a change is expected. A cyclical, iterative approach contributes to the improvement of existing practices (O'Leary, 2007).

The instruments for collecting the qualitative data included an interview with the lecturer about her experiences with the new interactive course design and a feedback survey from the students.

At the end of each block (section) the students wrote a test, and the results were recorded in the grade center of the LMS. These grades incorporated the quantitative data of the study.

Both sets of qualitative data were used to identify challenges, successes and suggestions. Additionally, the quantitative data were validated by comparing the grades of the 2022 and 2023 cohorts. The three sets of data were triangulated to support the findings.

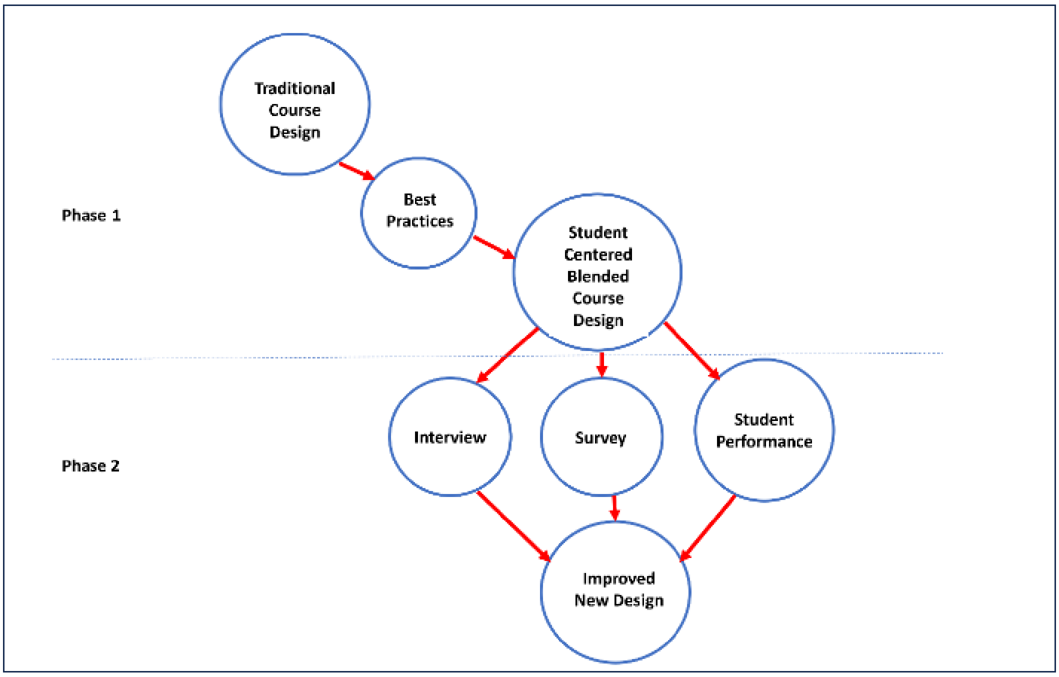

Figure 1: Course design phases

In this study, iterative action research was implemented over two phases. Phase 1, refers to redesign of the "old traditional course" based on best practices. In Phase 2, a student-centered course structure is designed and implemented over a year (referred to as "new" design). Thereafter, the lecturer reviews the old and new designs during an interview with the researcher. Additionally, the student survey provides the researcher with their opinion about on the new course design. The performance of the student from the "old course" and the "new course" are compared to find their opinions on the "old" and "new" course designs. These data sets provide valuable input for designing an improved course.

Phase 1 - Description of the old course

The course was built on the LMS of the University by the lecturer. Before the re-design of the old course the lecturer used the LMS for loading the course content and some quizzes without a specific structure.

The lecturer loaded resources such as PowerPoint slides and Lecture Notes for students to read after classes and summative assessments in the form of quizzes for revision of learning material to the course. These documents and quizzes were loaded in folders without a specific structure.

The "old" course was not used as an active teaching and learning platform but rather as a repository for the content followed by formative quizzes for testing the students' knowledge after a section. The lecturer followed a traditional lecturer-centered approach with limited student engagement.

The "old" course included two face-to-face lectures and one clinical practical session a week. The amount of information was enormous, and the lecturer followed a lecturer-centered approach to cover the load of content as fast as possible.

The timelines included the first phase (Fig 1) which was conducted at the end of 2022 when the old course was analysed.

The second phase (Fig 1) was done at the beginning of 2023 where additional activities were added based on Merrills Instructional design model.

After the changes were made and the course ran for a year in 2023, the lecturer and students were requested for their feedback. The feedback would assist the designer to improve the design based on action research principles where a study is repeated with improved iterations.

The "old" lecturer centered course, was analyzed based on the ADDIE instructional design model. Additionally, Merrill's principles of first instruction (2012), were used to improve student's engagement with the content. This model includes task centered, activation, demonstration, application and integration of new content (Merrill, 2012).

Phase 2 – New course design

In Phase 2, a student-centered course structure was designed and implemented over a year (referred to as "new" design). After considering best practices, the lecturer created learning modules which were divided into terms and in each learning module where all resources and activities for that term were included in respective folders.

After the changes were made the course ran for a year in 2023 and then the lecturer and students were requested for their feedback. The feedback would assist the designer to improve the design based on action research principles where a study is repeated with improved iterations.

The organization of the content was needed to improve navigation and to offer learning paths to accommodate learning preferences of students, with the student in mind.

Activities that were added to the new designed course, included:

- Term 1 -Video recordings by students on the upper thorax region as group work and uploaded in the discussion area of the Learner Management System,
- Term 2 - Students drawing and using clay to build muscles around the skeleton.
- Term 3 - Compiling question banks and quizzing each other.
- Term 4 - Lecturer compiling Quiz Questions with Kahoot;

The aim of the design focused on increased student engagement with the content.

5.1 Evaluation

After Phase 1 and 2 were implemented and then the re-designed "new" course was reviewed by the lecturer and the students based on the research questions of the study. These questions included the challenges, successes and suggestions of redesigning a Human Anatomy course to achieve a more student-centered approach and ultimately aiming at improving the student's engagement, enthusiasm and performance.

5.2 Sampling and Population

The study utilized purposive sampling. This approach is relevant for classes with student participants that are readily available (Kandola; 2014). The sampling population for this study included the instructor and a class of 60 students.

5.3 Reliability and Validity

The reliability and validity of the data analysis lies in the analysis of the data by searching for specific themes and by member checking with a research participant (the instructor in this case) to ensure that the researcher understood the interview questions and answers them correctly (Harvey; 2015). The following steps were taken to ensure validity and reliability of the data. Step 1: Becoming familiar with the data, Step 2: Generating initial codes, Step 3: Searching for themes, Step 4: Reviewing themes, Step 5: Validate themes by a research participant. Step 6. Defining themes, Step 7: Write-up. (Braun and Clarke; 2012).

5.4 Data Collection, Analysis and Results

During the second phase in Fig.1, the first set of data was collected during a virtual interview with the lecturer. Thereupon the transcript of the recording was downloaded and analysed.

The second set of data collection was retrieved from a post-course survey from participating students. And the third set of data was collected from the students' grades from the previous and the current courses which were retrieved from the learner management system. After the data collection the data was analysed as described below.

5.4.1 Data collection 1 – the interview

The data from the interview with the lecturer was analysed by applying an inductive thematic analysis (Williams; 2007). Where the data are read and scrutinized by searching for patterns in the data which are organized into themes. The researcher makes sense of the data from his/her perspectives by applying an active reflective process. The perspective is based on best practices in the field of instructional design (Braun and Clarke; 2021).

A "spreadsheet" software was used for data analysis. The identified themes were added to the spread sheet in verbatim. An inductive approach was followed. This is a flexible analysis approach which means that the researcher allows the data to guide the analysis. This approach is used to identify emerging patterns, themes and concepts (Williams; 2007). The researcher reflects on each theme in the context of the course. The themes were identified and compiled into Table 1.

5.4.2 Data collection 2 - survey

For the second data capture and analysis, the students were requested to give their feedback by completing a survey. After collecting the data an inductive thematic analysis was applied (Williams; 2007).

The survey questions were divided into challenges, successes and suggestions. The answers were sorted and analysed by using spread sheets again. Specific concepts were found under each theme. The themes and concepts are added to the last column of Appendix 1, (Table 3).

A thematic analysis was conducted to identify relevant themes in the lecturer's and students' feedback. By applying member checking the analysis and interpretation of the interview was validated by the research participant.

5.4.3 Data collection 3 - student grades

The third set of data was collected from the student performance during 2022 and 2023 (Table 1).

5.5 Data Analysis

The instruments for the data analysis include an interview, a survey and the performance of the students. All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of SMUREC (SMUREC/A/310/2023:IR).

5.5.1 Interview

The data from the interview with the lecturer was analysed by applying an inductive thematic analysis (Williams; 2007). Where the data are read and scrutinized by searching for patterns in the data which are organized into themes. The researcher makes sense of the data from his/her perspectives by applying an active reflective process. The perspective is based on best practices in the field of instructional design (Braun and Clarke; 2021).

A spreadsheet software was used for the data analysis. The identified themes were added to the spread sheet in verbatim. An inductive approach was followed. This is a flexible analysis approach which means that the researcher allows the data to guide the analysis. This approach is used to identify emerging patterns, themes and concepts (Williams; 2007). The researcher reflects on each theme in the context of the course. Themes were identified and presented in Table 3.

5.5.2 Student survey

For the second data capture and analysis, the students were requested to give their feedback by completing a survey. After collecting the data an inductive thematic analysis was applied (Williams; 2007).

The survey questions were divided into challenges, successes and suggestions. The themes and concepts were compiled in a table format and are added in the last column on Addendum 1. A thematic analysis was conducted to identify relevant themes in the lecturer's and students' feedback. By applying member checking the analysis and interpretation of the interview were validated by the research participant.

5.5.3 Student performance.

The third set of data was collected from the student performance of two cohorts. These included data from their final grade.

6. Bias and Limitations of the Study

The researcher may have been biased because her field of research is directly related to designing courses to improve student learning. This may influence her judgement as her positive approach may affect her search for positive factors that support her field of study. To reduce any bias, both the instructor and an external reader have critically reviewed the thematic analysis and conclusions.

7. Findings and Conclusion

The conclusion aims at answering the research questions which included the following: "What strategies can be implemented to improve student performance in a Human Anatomy course?"

The challenges, success and suggestions for the new course, as experienced by the participants, are presented in Table 2 with a more detailed explanation in Table 3 and 4.

Besides the importance of improved student engagement, the performance also indicates that the student grades improved. The previous cohort from 2022 indicated a high percentage of absenteeism and that students became too stressed and worried about the class, to enjoy it.

By adding best practices to phase 1 the updated course design improved. This was also supported by the improvement of the students' throughput rate. These results could be the breakthrough the relevant University had been hoping for because currently the high retention rate of the second year Human Anatomy cohort, is concerning as it blocks the students' progress at the university.

The comments from the interview and the surveys from the qualitative data are supported by the quantitative data. The grades indicate an increase of 10% on the class average compared to the class percentage of the previous cohort. Another astounding result is the decreased year's failure rate, from 30% to 16% (by 14%) as depicted in the Figures 2 and 3 below.

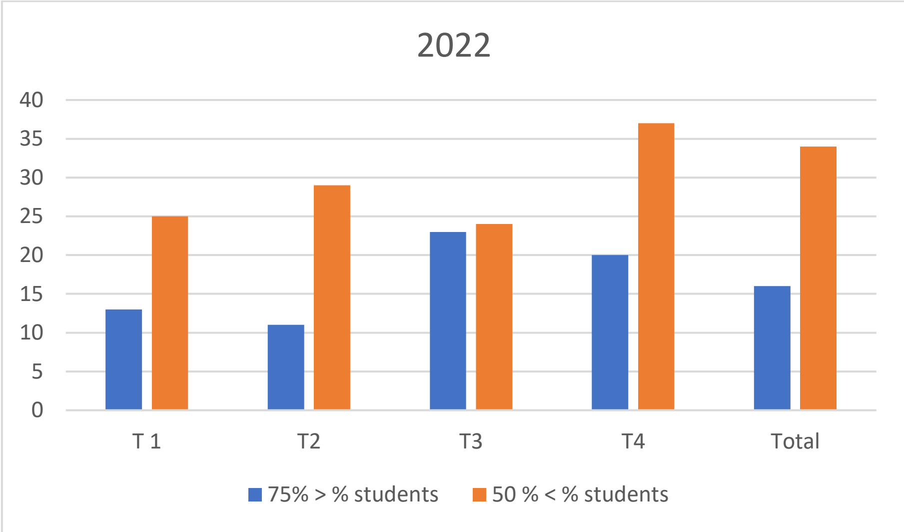

Figure 2: Student performance 2022

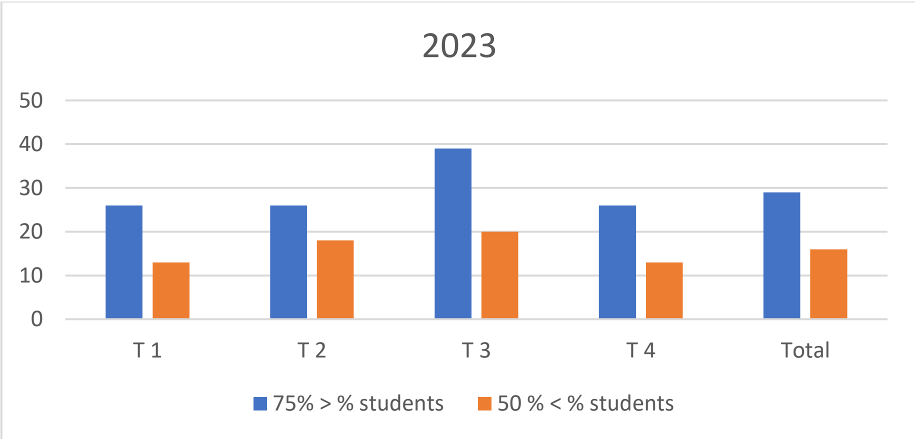

Figure 3: Student performance 2023

The most impressive result of the redesigned course was the increased student engagement with the course content and the high percentage of the students from the BB training class receiving distinctions. The distinction percentage increased from "16%" in the "old" course to "29%" in the "new" course, and thus the students achieved a class average of 75%, in the "new" course.

Besides the improved performance of the students the positive results were supported by the words of the lecturer who ended the interview with the comment: "and there were no troublemakers in the class this year". This statement supports the importance of the positive involvement of the students. The results also indicate the remarkable change from being anxious about failing behind in the class to being motivated and to achieving success. From these results a strategic course design framework based on the results in Table 1 can be proposed for future use.

Table 1: Comparison of Student performance 2022 and 2023 per term and year

	2022	T 1	T2	T3	T4	Total	2023	T 1	T 2	T 3	T 4	Total
Percentage of class achieving more than 75% (Distinctions)		13	11	23	20	16		26	26	39	26	29
Percentage of class achieving less than 50% (Failures)		25	29	24	37	34		13	18	20	13	16

In conclusion, this case-study contributed to encouraging the lecturer to successfully ride the technology wave by re-designing a course based on best practices. Thereby improving the students' performance and motivation and ultimately improving from a class where a third of the class were failing to a class with low failures and a third achieving distinctions is a mentionable achievement. Therefore, this case-study can be recommended for student centered blended learning courses designs in the future.

AI Statement: The authors state that Artificial Intelligence tools were not used in this study.

Ethics Statement: Ethical approvals have been obtained with precautions taken to ensure participants' informed consent and confidentiality.

References

Alhojailan, M. I., and Ibrahim, M. 2012. Thematic analysis: A critical review of its process and evaluation. *West east journal of social sciences*, *1*(1), 39-47.

Baldwin, S., Ching, Y. H., and Hsu, Y. C. 2018. Online course design in higher education: A review of national and statewide evaluation instruments. *TechTrends*, *62*, 46-57.

Braun, V., and Clarke, V. 2021. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research*, *21*(1), 37-47.

Brown, T., and Katz, B. 2011. Change by design. *Journal of product innovation management*, *28*(3), 381-383.

Creswell, J. W., and Creswell, J. D. 2017. *Research design: Qualitative, quantitative, and mixed methods approach*. Sage publications.

Dixson, D. D., and Worrell, F. C. 2016. Formative and summative assessment in the classroom. Theory into practice, 55(2), 153-159

Garcia, F. 2000, January. CACTUS: Automated tutorial course generation for software applications. In *Proceedings of the 5th international conference on Intelligent user interfaces* (pp. 113-120).

Groh, F. 2012. Gamification: State of the art definition and utilization. *Institute of Media Informatics Ulm University*, *39*, 31.

Harvey, L., 2015. Beyond member-checking: A dialogic approach to the research interview. *International Journal of Research & Method in Education*, *38*(1), pp.23-38.

Kandola, D., Banner, D., O'Keefe-McCarthy, S., and Jassal, D. (2014). Sampling Methods in Cardiovascular Nursing Research: An Overview. *Canadian Journal of Cardiovascular Nursing*, *24*(3).

Koehler, M. J., Mishra, P., Hershey, K., and Peruski, L. 2004. With a little help from your students: A new model for faculty development and online course design. *Journal of Technology and Teacher Education*, *12*(1), 25-55.

Mazur, E. 2013. Peer instruction. *A User's Manual. [Kindle Edition]. Retrieved from Amazon. com*.

McKim, C. 2023. Meaningful Member-Checking: A Structured Approach to Member-Checking. *American Journal of Qualitative Research*, 7(2), pp. 41-52.

Merrill, M. D. 2012. *First principles of instruction*. John Wiley and Sons.

Mondal, H., Swain, S. M., and Mondal, S. (2022). How to conduct descriptive statistics online: A brief hands-on guide for biomedical researchers. *Indian Journal of Vascular and Endovascular Surgery*, *9*(1), pp. 70-76.

National Research Council, 2002. *Learning and understanding: Improving advanced study of mathematics and science in US high schools*. National Academies Press.

O'Leary, Z., 2007. Action research. *The Social Science Jargon-Buster*.

Schwab, K. 2017. *The fourth industrial revolution*. Currency.

Tucker, B. (2012). The flipped classroom. *Education next*, *12*(1), 82-83.

Williams, C. 2007. Research methods. *Journal of Business and Economics Research (JBER)*, *5*(3).

Appendix 1: Tables 2, 3 and 4

Table 2: Data collection – Interview and survey

Lecturer's perspective (Interview)		
Challenges of old design	**Successes of new design**	**Suggestions for the next updated design**
Quizzes of learning only Limited group participation Time management and work ethics Low level learning Low engagement with the content Low motivation Technology challenges Low take up of pre and post-tests Low group work participation Increased grading time Group activities took time	Quizzes for and of learning Increased group participation Enjoyment of Gamification, Increased time for practical sessions Increased student engagement Increased competition Increased collaboration Increased creativity	Adding grades to pre- and post-tests and to their total year marks. Improving active participation in group work. Adding rubrics to improve understanding of criteria and decrease grading time.
Students' perspective (Survey)		
Challenges of old design	**Successes of new design**	**Suggestions for the next updated design**
Complicated content High workload Challenge with time management Limited time to complete assignments Too fast teaching pace Low audibility in class Tutors can't always assist High Anxiety	Students' confidence grew. The course was well organized, Excellent lecturer Better understanding Better guidance Increased knowledge checks Improved knowledge retention Increased motivation	Access to previous tests More time to spend with specimens Lectures should be chunked and not too hurried. 10-minute breaks between 40-minute sessions No morning classes All demonstrators need to know their work Tests need to be spaced out over two weeks. More moderators are needed in group activities.

Table 3: Data Analysis

Feedback on courses before design changes	
Challenges from the lecturer perspective	**Explanation**
• Limited revision	• Quizzes were only conducted as assessment of learning and not for learning.
• Limited group participation	• Not all students participated in group projects.
• Time management and work ethics	• The students were struggling with time management and work ethics including procrastinating, low completion of self-assessments, low activity in pre- and post-lecture quizzes, low preparedness of students). They were also struggling with understanding = the content and dealing with high workload

• Low level learning • • Low engagement with the content • Low motivation • Technology challenges • Low take up of pre and post-tests • Low group work participation • Group activities were new and intimidating at first • Group activities take time • Individual activities needed buy-in • Grading time for participant increased	• Their learning approach was at a low level because they mostly learned an evening before the test. • This resulted in low understanding of and low interaction with the content and finally low performance. • To complete tasks and in group work not everyone participated. • Wi-Fi challenges in lecture rooms and on campus as well as in student residences. • Students ignored the tests because they did not see the use of them. Participants stopped them. • Group work not everyone participates • Increased grading time for group projects • students would not participate if they do not know why.
Challenges from the students' perspective before the redesign of the course	**Explanation**
• Complicated content • High workload • Challenge with time management • Limited time to complete assignments • Too fast teaching pace • Low audibility in class • Tutors couldn't assist • High Anxiety	• They needed to grasp the content better • The workload was too high • The students had bad time management challenges • They had limited time to complete the assignments • The lecturing pace was too fast, and they were battling to keep up. • They could not hear in class • The tutors couldn't always help • They got anxiety for the course • They needed more quizzes and practice questions • They did not have access to all quizzes
After the course was re-designed the following **successes** were identified from the interview and confirmed by the survey data:	**Explanation**
• **Improved group work** • **Additional quizzes** • **Gamification**	• Increased face to face and online collaborate activities were included where students learned with and from each other. Both the lecturer and the students found them to be helpful to increase understanding and engagement with the course material. • Pre-, post and formative quizzes were included for knowledge checks before and after classes and to practicals to assess the students' learning and understanding as soon as possible. Mock tests also helped to reduce anxiety levels. • Games were mentioned by both the lecturer and students. The students enjoyed challenging each

• **Practical sessions**	other by competing with quiz games such as Kahoot. • "Practicals" were very helpful for learning and understanding the content. Apprehension to work with specimens became better with more practical work.
Additional successes mentioned by the lecturer	**Explanation**
• Increased student engagement- • Increased competition – • Increased collaboration - • Increased creativity- • Better understanding – • Better guidance – • Increased knowledge checks – • Improved knowledge retention - • • Using study guide for outcomes guidelines and for questions- • Improved rubric creation skills • Improved the course goal awareness - • Adding interactive VR visual tools (Primal pictures)- • Better time management – • Increased engagement with content • Improved engagement with content with added activities.	• Students became more engaged with the content. • Games challenged students to become more active. • Group activities contributed to students learning from each other. • Video recordings of dances to indicate muscle sections and using colour for learning diagrams stimulated the students' creative thinking • Demonstrating concepts and terminology to increase better understanding • Guiding students' activities with rubrics • Adding pre- and post-test. • Students creating quiz questions for increased engagement with content. • Aligning the course design with the study guide outcomes • Reducing grading time with a good rubric • Encouraging students to use their study guide as a guideline from the beginning of the year. • Improving student engagement with multi-media • Flexible assignments time for students to work from home • Variety to their usual classroom experience.
Successes mentioned by students	**Explanation**
• Their confidence grew. • The course was well organized • Excellent lecturer	• By increased knowledge • They knew exactly what was going on. • She is very **organized**
Suggestion from the lecturer for future course design includes the following points:	**Explanation**
• Adding grades to pre- and post-tests and to their total year marks. • Improving active participation in group work. • Adding rubrics	• It is crucial to add points to pre and post-tests for the class work as students do not consider taking tests without them counting towards their grades. • For inactive group members group leaders should only add the names of active participants to group work. • To cut down on grading time.
Suggestions from students	• **Explanation**
• Access to previous tests • Need more time to spend with specimens	• For preparation • More practical time is needed

• Lectures should be **chunked and not too hurried**. • **10-minute breaks** between 40-minute sessions • **No morning classes** • All **demonstrators need to know their work** • **Tests need to be spaced** out over two weeks. • **More tutors** are needed in group activities.	• Lecturers should work slower in class • They need Breaks during face-to-face classes. • They prefer afternoon classes. • All tutors should be present. • Tests should be taken every two weeks. • More support is needed during group activities.

Table 4: Summary of Findings

Lecturer's perspective		
Challenges of old design	**Successes of new design**	**Suggestions for the next updated design**
• Quizzes of learning only • Limited group participation • Time management and work ethics • Low level learning • Low engagement with the content • Low motivation • Technology challenges • Low take up of pre and post-tests • Low group work participation • Increased grading time • Group activities took time	• Quizzes for and of learning • Increased group participation • Enjoyment of Gamification, • Increased time for practical sessions • Increased student engagement • Increased competition • Increased collaboration • Increased creativity	• • Adding grades to pre- and post-tests and to their total year marks. • Improving active participation in group work. • Adding rubrics to improve understanding of criteria and decrease grading time.
Students' perspective		
Challenges of old design	**Successes of new design**	**Suggestions for the next updated design**
• Complicated content • High workload • Challenge with time management • Limited time to complete assignments • Too fast teaching pace • Low audibility in class • Tutors can't always assist • High Anxiety	• Students' confidence grew. • The course was well organized, • Excellent lecturer • Better understanding • Better guidance • Increased knowledge checks • Improved knowledge retention • Increased motivation	• Access to previous tests • More time to spend with specimens • Lectures should be chunked and not too hurried. • 10-minute breaks between 40-minute sessions • No morning classes • All demonstrators need to know their work • Tests need to be spaced out over two weeks. • More moderators are needed in group activities.

Slideshows as a Tool for Learning and Assessment: Pros and Cons as Perceived by Students

Nitza Davidovitch and Rivka Wadmany
Education Studies, Ariel University, Israel

d.nitza@ariel.ac.il
rivkaw@ariel.ac.il

https://doi.org/10.34190/ejel.23.2.3647

Abstract: In academic studies, some course assignments take the form of presentations. The art of presentation involves conveying messages and one of the methods is by delivering presentations, either face-to-face, synchronously, and/or asynchronously. Presentations require analyzing a topic, processing an article, analyzing ideas, dilemmas, lesson plans, etc. The current study examines the benefits and shortcomings of slideshow presentations as a means for learning and as an assessment tool, as perceived by students. The study included 66 respondents, undergraduate students in a department of education at an Israeli university. Slightly more than half were studying education, 24.2% were studying education and for a teacher diploma, and 24.2% were studying for a teacher diploma. The sample included 13 men, who constituted 20% of the entire sample, and 53 women who constituted 80% of the sample. The mean age was 28. The study combined qualitative and quantitative tools. From the research findings it is evident that the students perceive the benefits of slideshow presentations as improving learning capabilities, as well as enhancing learning and the student's personal development. Another benefit emerging from the findings is that students see presentations as a means not only for presenting their knowledge to the lecturer but rather also for conveying information to their peers. Namely, the presentation has meaning both on the academic level and on the collaborative level. This benefit is perceived by the students as very important in the current era, where the collaborative dimension is receiving extensive place and emphasis in processes of teaching, learning, and assessment, as well as constituting part of students' learning outcomes. Creating a supportive learning community that generates a social-academic climate and boosts one's sense of efficacy and resilience in an era of change is meaningful and will remain with them in their future jobs as well. Concerning the difficulties experienced by students, they noted that slideshow presentations give rise to stage fright, a finding that tops the list of human fears. At the same time, the students noted as benefit the fact that facing an audience is an important skill that should be mastered, where through use of presentations they can attempt to overcome their inner difficulties and develop. The research findings shed light on the use of slideshow presentations as a tool and as a means for supporting learning and assessment. These findings may contribute to improving students' learning processes and to assessing their scholastic and personal development. For this purpose, however, it is necessary to train students to experience the preparation of effective presentations and to present to an audience, as well as to develop a culture that entails a collaborative-learning climate suitable for an era of digital innovation. Delivering presentations and speaking to an audience are essential skills in the modern workplace. Students will find themselves in professional situations that require them to present information, usually via slideshows and by conveying messages to an audience of listeners or viewers. It is possible to learn how to present to an audience and to acquire presentation skills, and there are even ways of successfully dealing with fear and enhancing confidence in such situations. Good strategies and techniques will grant the necessary training and tools to prepare and deliver presentations in an efficient and relevant manner.

Keywords: Slideshows, Learning, Assessment, Students, Academia

1. Introduction

Many academic faculty members integrate slideshows in their lectures to improve the teaching process and increase participants' interest (Bar, 2010; Salant, 2022). Slideshows are the most common tool for delivering lectures by faculty members. Students have varied attitudes to slideshow presentations: they see lecturers' presentations as a type of summary of the study material, a summary of the main points raised in class, while some see them as a means of illustrating the topics studied – depending on the nature and structure of the presentation, how it is utilized by the lecturer, and the coordinated expectations between the lecturer and the students (Grieve et al., 2021).

Intensive use of presentations arouses criticism and leads to many questions, such as: Is the teaching and learning process more interesting and attractive when using presentations? Can presentations promote creative thinking, or do they limit the space for thought? Do digital presentations necessarily lead to superficial delivering of the material? Are the ideas indeed presented in full, considering the need to reduce, focus, discern the main points?

Cite this article: Davidovitch, N. and Wadmany, R. 2025. "Slideshows as a Tool for Learning and Assessment: Pros and Cons as Perceived by Students", *Electronic Journal of e-Learning*, 23(2), pp 30-45, https://doi.org/10.34190/ejel.23.2.3647

This study examines students' perceptions of presentations and slideshows concerning the benefits and shortcomings involved in use of this tool for processes of learning and assessment. The research questions include the following: Do presentations contribute significantly to students' learning process? What are the strengths and weaknesses of lecturers' use of presentations? Moreover, lecturers tend to assess students in their courses based on the work they do, how they portray this work using slideshows, and the presentation of their work. What do students think of this manner of assessment?

2. Literature Review

2.1 Using Slideshow Presentations in the Classroom

In slideshow presentations a presenter displays a certain topic to an audience. Over the years, the research literature has related to the benefits and shortcomings of slideshow presentations used as a tool for conveying information (Salant, 2022). Several benefits of slideshows as presentation tools can be mentioned (Hammer & Ankori, 2008; Bar, 2010):

Potential benefits of integrating slideshows in class

- **Enhancing the lecture's clarity:** Helping the user organize information systematically and clearly, using titles and subtitles.
- **Saving time:** Use of slideshows saves time for class management and allows the presentation of graphs, tables, and pictures (Hadiyanti & Widya, 2018).
- **Facilitating interest and curiosity**: Slideshows can change the nature of the lesson and create a richer language comprised of symbols that make it possible to vary how messages are conveyed (Osman et al., 2022).
- **Processing information**: Students are exposed to ideas that are clearly worded and formulated. They can copy the lesson plan and do not need to concentrate for long to process the information (Widhayanti & Abduh, 2021).
- **Presenting the lecturer's "road map"**: The slideshow can help lecturers remember the lecture's structure and manage the lesson.
- **Maintaining efficient class management**: Slideshows reduce lecturers' need to dictate the contents of the lesson.

Other studies indicate some **shortcomings of using slideshows as a presentation tool** (Kozminsky, Gad, & Baraki, 2000; Mercer, 1996, Fisk, 2019):

- **Compromised ability to concentrate in class**: When the lecturer focuses on the slideshow without introducing changes in class or without diversifying the teaching methods. Continuous use of slideshows throughout the entire lesson, even if they are well designed, diminishes listeners' ability to concentrate. At times, **excessive auditory and visual effects** overload the lecture; also, some lecturers make unnecessary use of bright colors and different fonts, though these do not serve the content of the lecture and only distract the students (Meishar-Tal, 2011; Salant, 2022).
- **Dictating the structure of the lesson**: pre-planned slideshows may limit the dynamic flow of the lesson following students' responses and restrict participants' space for thinking during the lesson. Moreover, lecturers who wish to display the "road map" of their lecture and to present the entire lesson plan at the beginning detract from students' inquisitiveness and interest in the lesson.
- **Deficient internalization of the study material:** Quite a few lecturers use slideshows inadequately. They read the text out loud without adding explanations and display overloaded slides that make it hard to understand the study material. Students are not required to try and organize the information in their own words. In contrast, in lessons with no slideshow they must listen to the lecturer and summarize the information. When students take notes during a lecture, this facilitates information processing. Their cognitive level must include comprehension and not only accumulating information or recalling. They formulate the ideas in their own words (Salant, 2022).
- **Presenting a partial picture:** Slideshows have a built-in failure; one of their constraints is the use of short messages or main points instead of thorough and expansive analysis of ideas, so much so that they form a superficial impression. Slideshows might compel users to rely on points to present contents, whereby ideas are shortened into mere titles. As a result, students are exposed to a partial and distorted result. Notably, some lecturers have "liberated themselves from the shackles of the slideshow"; they present a single slide and talk about it for a lengthy time. Hence, the slideshow might

radically reduce the presentation of data. On the other hand, an overloaded slideshow has the effect of overburdening the students (Meishar-Tal, 2011).

- **Lack of contact with students:** Lecturers who maintain lengthy eye contact with the computer screen may reduce their eye contact with the students.

In a meta-analysis of 40 studies, Christof Wecker (2013) found that slideshows result in superficial learning processes and that they serve more as an aid for lecturers than for students. In his view, good lecturers are those who ask questions and arouse interest and not necessarily those who utilize slideshows. The spoken word is much more effective than a visual slideshow and students' tendency to concentrate on visual means may detract from the cognitive deciphering of the written word. Research findings show that fascinating lessons are not associated automatically with slideshows but rather with how lecturers activate the students – and this has many possible avenues, including how slideshows are utilized (Salant, 2022).

Most studies conducted to date on slideshows have disregarded the association between students' cognitive style, slideshows, and the learning process (Margaret, 2015). Studies show that to enhance the connection between students' cognitive style, slideshows, and the learning process it is necessary to strengthen the pedagogical aspects of the slideshow and focus on visual elements. In most cases, the lecturers are those who plan and prepare the slideshow, while also receiving a more thorough grasp of the study material (Salant, 2022). One recommendation is to change direction and allow the students themselves to plan and prepare the slideshows, such that they will be more deeply acquainted with the study subject or research subject. Recently, we are discovering that lecturers are utilizing this constructivist approach and charging their students with preparing and presenting collaborative slideshows (Meishar-Tal, 2011; Kozminsky, 2000; Margaret, 2015).

In addition to the studies already reviewed regarding the benefits and drawbacks of slideshows as a presentation tool, recent research highlights the importance of effective presentation techniques that enhance both learning outcomes and students' presentation skills. This is particularly relevant as presentations have become a common learning and assessment tool in higher education. Moreno and Mayer (2007) conducted a large-scale quantitative study examining the effectiveness of multimedia presentations that combined short texts, images, and animations, compared to text-only presentations. Their findings demonstrated that students who were exposed to multimedia-enriched presentations achieved significantly better comprehension and retention than those who experienced traditional text-based lectures. Grieve (2021) quantitatively assessed the impact of body language, varied intonation, and eye contact on audience engagement during academic presentations. This study surveyed over 200 presentations at academic conferences and revealed a strong positive correlation between effective delivery techniques and audience attention and involvement.

Lee, Kang and Park (2023) conducted an experimental study to evaluate the effect of incorporating personal storytelling into academic presentations. Their results indicated that storytelling enhanced both students' emotional engagement and their ability to recall information, suggesting that narrative techniques foster deeper cognitive and emotional processing.

Gallo (2014) conducted a qualitative study based on in-depth interviews with leading academic lecturers, exploring their preferred presentation techniques. The analysis showed that techniques such as opening with a personal anecdote, using light humor, and actively engaging the audience through dialogue were identified as particularly effective for maintaining attention and fostering interaction. Carmine (2018) carried out qualitative content analysis on 50 highly successful TED talks, identifying patterns of presentation techniques that contributed to their success. Key findings highlighted the consistent use of powerful visuals, memorable soundbites, and emotional appeals as core strategies for impactful communication. Jones and Sheppard (2016) conducted an ethnographic study within communication courses at the university level, examining how students develop presentation skills. They found that the most effective learning occurred through repeated simulations, peer feedback, and reflective learning, underscoring the importance of iterative practice and collaborative learning environments.

Taken together, these studies emphasize that successful academic presentations require more than well-designed slides; they depend on the presenter's ability to use their voice, body language, narrative structure, and audience interaction effectively. This reinforces the argument that preparing students for high-quality presentations requires training in both content development and delivery techniques, supported by clear guidelines and constructive feedback throughout the learning process.

2.2 Slideshows and Presentations as Assessment Tools

In recent years, an era when lecturers are required to teach using new technological tools, an era of artificial intelligence (AI), an era of teaching and learning in situations of uncertainty and change such as Covid-19 – slideshows created by students are becoming a tool for assessing their performance in the course. Different approaches to assessment discern between summative and formative assessment. The traditional, prevalent type of assessment is based on quantitative elements, where the object assessed is usually the product (Jacobson & Spiro, 1995). This type of assessment is carried out at the end of a unit/chapter and constitutes an endpoint for receiving information on the student's achievements or for reaching judgments and decisions concerning students' placement in a class/group. In contrast, formative assessment relates both to the learning process and to the learning products and is carried out at all learning stages; it is intended to provide the teacher and students with feedback on progress in the study material and on specific difficulties. Formative assessment is usually descriptive, detailed, and makes it possible to address non-quantifiable aspects such as originality, spontaneity, creativity, teamwork, and intensive analysis (Meishar-Tal, 2011).

Research findings show that students of education ascribe a great deal of significance to producing slideshows for purposes of learning and some prefer to produce a slideshow over writing a final paper (Meishar-Tal, 2011; Kozminsky, 2000; Margaret, 2015). These studies contend that exams are usually given once or twice a semester, while when producing a slideshow, the assessment is procedural, collaborative, and occurs several times: self-evaluation, peer evaluation, evaluation by the teacher, and others. It is possible to participate, update, explain, persuade.

The study performed by Meishar-Tal (2011) shows that most of the potential criteria and unique features noted were found significant by different groups of learners. Nonetheless, the criteria of reflectivity and the criteria of the conscious significance of the learning discourse have been emphasized relatively less than other potential criteria. Regular classes take exams once or twice a semester, while when producing a slideshow, the assessment is procedural, collaborative, and occurs several times: as self-evaluation, peer evaluation, evaluation by the teacher, and others. But it is possible to participate, explain, convince.

The research findings show that reservations have been voiced regarding including students in determining the assessment, for the following reasons: assessment that is mostly summative and is manifested in a grade is perceived by students as formal, comparative, sometimes threatening, and therefore it is important that it be carried out by the teachers. The teachers see it as part of their job, a product of their accumulated experience, and part of their professional authority. Teachers have the competence and experience to award a final grade.

Some have noted that a learning process that includes producing a slideshow involves different assessment methods than those customary in traditional classes. The process is important, and it suits experienced teachers.

In the modern era the use of slideshows as a presentation tool has become routine. Presentation is an important acquired skill that improves with practice. Well-made slideshows allow good portraying of information, arousing thought, moving, convincing, and even mobilizing the viewers to action. Salant (2022) relates to ways of integrating slideshows in teaching, learning, and assessment, and recommends accompanying slideshow presentations with portfolios. In the portfolio students document each stage of their work process, and at the end of the assignment both the portfolio and the presentation are assessed.

Lecturers who use this method of assessment perceive presentation as an acquired skill that improves with practice, and it is part of developing collaborative assessment. Some see assessment as a continuous multi-phase dialogue that involves continuous interaction between the teacher and students and among the students themselves. This discussion is at the heart of the constructivist approach to learning as a social process. Alternative assessment (particularly via digital portfolios or performance-based assessment) makes it possible to relate to students' capabilities as manifested in their products (Tannenbaum, 1996). The current study examines the benefits and shortcomings of slideshow presentations as a means of learning and an assessment tool, as perceived by students. The research questions include the following: Does the presentation contribute significantly to students' learning process? What are the strengths and weaknesses of lecturers' use of presentations? Considering that lecturers tend to assess students in their courses based on the work they do, how they portray this work using slideshows, and the presentation of their work, what do students think of this manner of assessment?

3. Research Findings

3.1 Quantitative Analyses

The study consisted of 66 respondents, undergraduate students at the Department of Education in Ariel University in Israel. Slightly more than half were studying education, 24.2% were studying both education and for a teacher diploma, and 24.2% were studying for a teacher diploma. The sample included 13 men who constituted 19.7% of the entire sample, and 53 women who constituted more than half the sample. The mean age was 28, with a standard deviation of about 9 years. Most of the respondents were single and a small proportion were divorced. The reliability of measures for assessing the quality of teaching using presentations as an assessment tool. See Table 1.

Table 1: Reliability of measures for assessing the quality of teaching using presentations as an assessment tool

Assessment measures	Scale reliability (Cronbach's alpha)
Improving study capabilities	α = 0.93
Comfort	α = 0.74
Lecturer availability	α = 0.78
Improving teaching-interest	α = 0.88
Improving teaching-order and organization	α = 0.81
Improving teaching-clarity	α = 0.79
Innovation, creativity	α = 0.85
Personal preference (for presentations)	α = 0.91
Interpersonal interaction	α = 0.77

Measures for assessing the quality of teaching using presentations as an assessment tool were ranked on a Likert scale from 1 to 5, where a higher score indicates high assessment of the quality of teaching. Table 2 summarizes the means of measures for assessing the quality of teaching using presentations for the research participants.

Table 2: Means and standard deviations of measures for assessing the quality of teaching using presentations (N=66)

Improving study capabilities:	Mean	3.24
	Standard deviation	1
	Minimum	1
	Maximum	4.87
Comfort:	Mean	3.17
	Standard deviation	0.90
	Minimum	1
	Maximum	4.67
Lecturer availability:	Mean	2.70
	Standard deviation	0.98
	Minimum	1
	Maximum	5
Improving teaching - interest:	Mean	3.40
	Standard deviation	1.09
	Minimum	1
	Maximum	5
Improving teaching - order and organization:	Mean	3.17
	Standard deviation	1.15
	Minimum	1

	Maximum	5
Improving teaching - clarity:	Mean	3.28
	Standard deviation	1.22
	Minimum	1
	Maximum	5
Innovation, creativity:	Mean	3.33
	Standard deviation	1.25
	Minimum	1
	Maximum	5
Personal preference (for presentations):	Mean	3.34
	Standard deviation	1.15
	Minimum	1
	Maximum	5
Interpersonal interaction:	Mean	2.99
	Standard deviation	0.86
	Minimum	1
	Maximum	4.71

The assessment measure with the highest mean, according to the data, was "improving interest in teaching", with a mean of 3.40, followed by the measure of creativity, with a mean of 3.33. Preference for presentations as an assessment tool was also found to have a relatively high mean, at 3.34.

The association between self-evaluation variables, demographic variables, and variables concerning the efficacy of teaching using presentations as an assessment tool

To explore the association between variables related to self-evaluation of academic achievements, demographic variables, and variables concerning the efficacy of teaching using presentations as an assessment tool, a Pearson test was conducted. Table 3 presents a summary of the findings.

Table 3: Pearson correlation between demographic variables, self-evaluation, and variables concerning the efficacy of teaching using presentations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Age	--														
2	Gender female	-.168	--													
3	Grade average	.337**	.087	--												
4	Self-evaluation as student	.210	.129	.663**	--											
5	Preference for face-to-face courses	.236	-.287*	.276*	.252*	--										
6	Preference for online courses	-.028	-.057	-.063	-.257*	-.167	--									
7	Preference for hybrid courses	-.101	-.125	.032	.007	.207	.539**	--								
8	Improving study capabilities	.178	-.110	.385**	.316**	.681**	.063	.331**	--							
9	Comfort with face-to-face presentations	.189	-.184	.457**	.484**	.704**	-.393**	-.026	.602**	--						
10	Lecturer availability	.143	-.270*	.113	.150	.345**	-.031	.169	.487**	.349**	--					
11	Improving teaching-interest	.236	-.197	.321**	.311*	.640**	.066	.362**	.876**	.524**	.466**	--				

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
12	Improving teaching-order and organization	.287*	-.094	.334**	.200	.524**	.165	.362**	.706**	.446**	.474**	.748**	--			
13	Improving teaching-clarity	.187	.036	.331**	.276*	.519**	.038	.255*	.870**	.463**	.332**	.788**	.682**	--		
14	Innovation/creativity	.160	-.005	.269*	.269*	.510**	-.029	.284*	.808**	.541**	.443**	.754**	.703**	.750**	--	
15	Personal preference for presentations	.173	-.066	.333**	.326**	.685**	.067	.355**	.900**	.624**	.393**	.869**	.712**	.815**	.773**	--
16	Interpersonal interaction	.213	-.020	.273*	.237	.577**	.073	.346**	.868**	.475**	.592**	.775**	.678**	.750**	.749**	.755**

*Correlation is significant at the 0.05 level (2-tailed); **Correlation is significant at the 0.01 level (2-tailed).

According to the findings, a strong positive association is evident between the student's self-evaluation of his or her achievements and their sense of comfort with face-to-face presentations in class. A positive association with medium power was also found between the perception of presentations as improving study capabilities and reporting a high average grade in the course. Other positive associations, although with lower power, were found between the student's self-evaluation of themselves and their achievements and the perception of presentations as contributing to interest, order and organization, clarity of the material, and improving interpersonal interactions in the course. A weak negative association was found between gender-female and preference for face-to-face courses, as well as lecturer availability. A weak negative association was also found between self-evaluation of the student's achievements and preference for presentations in online courses.

The association between attendance levels in courses where students delivered presentations and variables comprising the efficacy of teaching using presentations as an assessment tool

To explore the association between the level of attendance in courses where students gave presentations and variables comprising the efficacy of teaching, a Pearson test was conducted. Table 4 presents a summary of the correlation's measures.

Table 4: Pearson correlations between level of attendance in presentation classes and variables comprising the efficacy of teaching using presentations

		1	2	3	4	5	6	7	8	9	10	11
1	Low attendance	--										
2	Moderate attendance	-.252*	--									
3	High to very high attendance	-.377**	-.801**	--								
4	Improved study capabilities	-.276*	-.103	.269*	--							
5	Comfort with face-to-face presentations	-.259*	-.258*	.407**	.602**	--						
6	Lecturer availability	-.013	-.147	.148	.487**	.349**	--					
7	Improved teaching-interest	-.158	-.119	.212	.876**	.524**	.466**	--				
8	Improved teaching-order and organization	-.065	-.236	.266*	.706**	.446**	.474**	.748**	--			
9	Improved teaching-clarity	-.325**	-.144	.338**	.870**	.463**	.332**	.788**	.682**	--		
10	Innovation/creativity	-.251*	-.030	.184	.808**	.541**	.443**	.754**	.703**	.750**	--	
11	Personal preference for presentations	-.251*	-.112	.262*	.900**	.624**	.393**	.869**	.712**	.815**	.773**	--
12	Interpersonal interaction	-.209	-.085	.210	.868**	.475**	.592**	.775**	.678**	.750**	.749**	.755**

*Correlation is significant at the 0.05 level (2-tailed); **Correlation is significant at the 0.01 level (2-tailed).

The findings indicate a significant positive association between high to very high attendance of classes and comfort with face-to-face presentations. A positive association was also found between high attendance and the perception of presentations as contributing to clear understanding of the study material. Other positive associations, although with weak power, were found between high attendance of classes and improved study capabilities, order and organization, and personal preference for presentations as an assessment tool in the course. In contrast, students who reported low and/or moderate attendance of presentation classes were found to have a negative association with the measures of teaching efficacy: improving study capabilities, comfort with face-to-face presentations, improved teaching-clarity, innovation/creativity, and personal preference for presentations. The strongest negative association was found between low attendance and the perception of presentations as improving clear understanding of the study material. In general, the findings show that high to very high attendance indicates high values in measures of the efficacy of teaching using presentations, while low to moderate attendance indicates low values for the measures of teaching efficacy.

The association between level of experience with using presentations, types of tools used when giving presentations, and variables comprising the efficacy of teaching using presentations as an assessment tool

To explore the association between students' level of experience with giving presentations (number of times the student gave a presentation to a class), the types of tools used by students in presentations (videoclips, animation, pictures, links to information sources, interviews), and variables comprising the efficacy of teaching using presentations as an assessment tool, a Pearson test was conducted. Table 5 presents a summary of the correlation's measures.

Table 5: Pearson correlation between experience with giving presentations, types of tools used for presentations, and variables comprising the efficacy of teaching via presentations

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Number of presentations	--														
2	Use of texts	-.109	--													
3	Use of videoclips	.224	-.028	--												
4	Use of animation	.167	-.183	.232	--											
5	Use of pictures	**.319****	.072	.406**	.249*	--										
6	Use of links to information sources	**.319****	.190	.342**	.168	.422**	--									
7	Use of interviews	.062	-.036	.312*	.120	-.048	.107	--								
8	Improved study capabilities	.026	-.006	.163	.183	.025	.001	.154	--							
9	Comfort with face-to-face presentations	.066	-.195	-.040	.164	.016	-.157	-.079	.602**	--						
10	Lecturer availability	-.022	.148	.199	.028	.075	.135	-.030	.487**	.349**	--					
11	Improved teaching-interest	.035	-.036	.117	.241	-.017	.069	.018	.876**	.524**	.466**	--				
12	Improved teaching-order and organization	.013	-.069	.132	.160	-.010	.044	.002	.706**	.446**	.474**	.748**	--			
13	Improved teaching-clarity	.115	-.063	**.273***	**.250***	.182	.068	.123	.870**	.463**	.332**	.788**	.682**	--		
14	Innovation/creativity	.173	.012	**.285***	**.291***	.242	**.259***	.059	.808**	.541**	.443**	.754**	.703**	.750**	--	
15	Personal preference for presentations	.134	-.087	.134	**.308***	.101	.084	.107	.900**	.624**	.393**	.869**	.712**	.815**	.773**	--
16	Interpersonal interaction	.016	-.005	**.246***	.205	.074	-.016	.145	.868**	.475**	.592**	.775**	.678**	.750**	.749**	.755**

*Correlation is significant at the 0.05 level (2-tailed); **Correlation is significant at the 0.01 level (2-tailed).

According to the findings, a moderate positive association was found between experience with using presentations and use of pictures and links to information sources in the presentation, a moderate positive association between use of animation in presentations and personal preference for presentations. Other positive associations, albeit with low power, were found between use of video clips and use of animation, and the perception of presentations as improving teaching clarity and as improving innovation and creativity in teaching. Another positive association with weak power was found between the use of videoclips in presentations and perceiving the presentation as improving interpersonal interaction.

3.2 Open-Ended Questions

Following the quantitative analysis, the participants were asked to answer seven open-ended questions. In the first question the participants were asked to describe in their own words "What are the benefits of the assignment you presented by means of a slideshow?". All the students (100%) answered this question, where the large majority gave positive feedback on the presentation as an assessment tool. Only about 11% noted that there is no benefit to presenting using a slideshow, with no further details.

Approximately 36% of the participants noted that the presentation improves their study capabilities. For example, a student noted that through the presentation he managed to reach a more thorough understanding of the material and to better remember the study material: *"You study the topic better from all kinds of directions that are convenient for you in order to pass it on and this causes you to better remember the topic studied".* Other students noted that both the process of preparing the slideshow, and its presentation gave them a more thorough understanding of the study material. For instance: *"Increased knowledge of the topic", "I experienced the material actively, meaning that I created a slideshow and was able to present it, such that the knowledge became clearer".* About 35% of the students also noted that the presentation contributed to deeper learning, which led to a more thorough and clear grasp of the material: *"Deep and meaningful learning of the material", "allows a clearer grasp of the material studied and makes it possible to bring creative ways of conveying the knowledge "*, and also learning while experiencing: *"learning the material while exploring and experiencing and not by rote".*

Other students noted that beyond the more meaningful learning, the presentation helped them improve technical capabilities such as facing an audience and strengthened their self-confidence: *"overcoming concerns and reinforcing self-confidence", "carrying out the assignment requires handling the class, and this is beyond preparation of the slideshow".* The presentation, according to the participants, also contributed to their creativity and innovation: *"The possibility of improvising, using diverse means, answering questions that direct you to other places you did not address", "realizing creativity".*

About 20% of the participants noted that the interaction with the lecturer and their peers through the presentation was a positive component of the learning process, for example an improvement as part of the discourse between the presenters and the students and lecturer during the presentation: *"Forming eye contact with the lecturer, the possibility of explaining yourself immediately if something is unclear", , "that the lecturer heard me and didn't only read what I had written. She managed to understand me more clearly and precisely. As one who writes slowly and has attention disabilities, it was easier and better for me to speak than to write or consider what answer to opt for", "collaborative work in a group, oral expression skills".*

Other students reported additional benefits of giving presentations for understanding the study material, such as interest, order and organization, proving capabilities, etc. (Table 6).

Table 6: Analyzing themes - first question (n=66)

Themes	%	n
Improving learning capabilities	39	25
Improving learning capabilities-clarity	35	23
Improving learning capabilities-self-expression	33.3	22
Interaction with lecturers and students	20	13
Improving self-confidence	15.2	10
Proving capabilities	13.6	9
Improving learning capabilities-creativity	13.6	9
Active learning based on experiencing	13.6	9
Improving learning capabilities-innovation	7.6	5
Additional improvements (saving time, comfort, order and organization, interest)	16.5	11
No benefits	10.6	7

It is evident from the analysis that presentations as a method of assessment and as part of the learning process generate many significant benefits for students. Presentations contribute to more thorough understanding of the study material, better retention, and deeper and more meaningful learning. In addition, presentations help develop important skills such as facing an audience, self-confidence, creativity, and innovation. Hence, students who saw a significant benefit to use of presentations as part of the teaching and assessment process perceived it as potentially enriching their learning experience and leading to better learning outcomes.

In the second question, students were asked to describe the difficulties they encountered when preparing the slideshow or presenting it. All participants but one answered this question (98.5%). The most conspicuous theme was the concern and difficulty involved in facing an audience while making the presentation, both technically and emotionally. On the emotional dimension, for instance, respondents reported low self-confidence: "concern of being judged by other people, concern of a blackout, trembling voice during the presentation, fast respiration that makes it hard to breathe", "a concern of standing and talking before the class". Students also indicated concerns regarding their ability to express themselves concisely while facing an audience for a set time "Difficulty summarizing the text in the slideshow and remembering where to expand. Also, it was very difficult to face an audience of students and talk to them", "…the only difficulty was giving the presentation in a very short time span and summarizing everything. I felt that I didn't really manage to convey everything I had wished".

A considerable portion of the respondents noted a difficulty with condensing and clarifying the material also while preparing the slideshow, namely, a difficulty distinguishing between the significant and the insignificant and with proper presentation of the main points: "fine-tuning the text in the slides", "when preparing the presentation, I battled with the complexity of encompassing the wide topic and conveying the precise message in a limited time span". Also reported were difficulties with oral expression: "difficulty learning to present things by rote", "knowing what to say beyond what is written", as well as keeping to the schedule: "with the amount of knowledge that has to be included in the few minutes of the presentation".

Other difficulties noted, regarding the process of preparing the presentation, were lack of technological knowledge on preparing slideshows: "a few technological problems with preparing the slideshow". Students also reported difficulties involving lack of proficiency in the study material: "Analyzing and understanding the material accurately requires lots of energy", unclear instructions for the assignment: "misunderstanding the instructions", as well as lack of accompaniment and support by the lecturer: "that there is no focused guidance by the lecturers on how to perform assignments", "collaboration with peers; lack of time by the lecturer". Also, assessment that does not reflect one's performance: "stage fright and the feeling that the grade usually does not reflect the [actual] presentation".

In summary, the analysis shows that the major challenges of students with preparing a slideshow and presenting it stem from concerns and difficulties with facing an audience, trouble condensing and summing up the material, difficulties with oral expression and keeping to a schedule, and also a dearth of technological knowledge, lack of proficiency in the study material, unclear instructions for the assignment, and an absence of sufficient accompaniment and support by the lecturers (Table 7).

Table 7: Analysis of themes - second question (n=65)

Themes	%	n
Difficulties with giving a presentation to an audience	38	25
Difficulty pinpointing and condensing the material	19.7	13
Difficulties with free speech and retaining in memory	12.1	8
Difficulties with expression – conveying messages	12.1	8
Technological difficulties	10.6	7
Lack of proficiency in preparing a slideshow	9.1	6
Lack of time	7.6	5
Difficulties with teamwork while preparing the slideshow	7.6	5
Concern of being assessed	6.1	4
Lack of self-confidence	6.1	4
Difficulties with the lecturer	6.1	4
Misunderstanding the instructions	6.1	4
Lack of proficiency in the study material	6.1	4
No difficulties	12.1	8

In the third question, the students were asked: "Did you feel empowered by giving the presentation?". All participants but one answered the question (98.5%). The large majority gave short answers of yes (45.5%), no (39.4%), or partially (9.1%), with no additional details. A small proportion of the respondents (4.5%) reported a sense of relief immediately upon concluding the presentation (Table 8).

Table 8: Analysis of themes - third question (n=65)

Themes	%	n
Yes	45.5	30
No	39.4	26
Sense of relief upon concluding the presentation	9.1	6

In the fourth question, the students were asked: "Did you feel that your ability to convey messages to the group of learners was empowered by use of a presentation?". All participants but one answered the question (98.5%). Slightly more than half the respondents answered "yes" (51.5%), with no further details, and 38.9% answered "no" with no further details. Approximately 15% of the respondents also reported other positive sensations after giving the presentation, such as satisfaction with their self-expression capabilities. *"It was more that I was able to express my understanding and present it as I wished, rather than empowerment,* as well as sensations of enjoyment, capability, and confidence. *"Yes, the presentations gave me a sense of confidence and strengthened my capabilities. In addition, there were several courses where I enjoyed giving the presentations and experiencing an array of fields" (*Table 9).

Table 9: Analyzing themes - fourth question (n=65)

Themes	%	n
Yes	51.5	34
No	37.9	25
Positive feelings after concluding the presentation	15.2	10

In the fifth question, the students were asked to express their opinion about ways of improving the preparation and presentation of the slideshow: "What do you think should be done to improve the preparations for planning and presenting the slideshow?". Sixty-four of the 66 respondents answered this question (approximately 97%). As evident from their replies, some 38% of the respondents contend that clearer and more focused instructions should be emphasized: *"clearer instructions regarding the amount of text"*, as well as practical guidance by the lecturer concerning the material: *"giving detailed instructions on how to prepare it", "meeting with the lecturer before"*, or giving examples: *"giving more detailed explanations with prepared examples".*

Approximately 24% of the respondents noted that it is necessary to add workshops on general topics related to providing tools for proper presentation: *"Perhaps in class to give more preparation and tools for how to do it right and how to build a slideshow and not only instructions on the assignment itself", ,* giving tips for facing an audience: *"teaching tips for facing an audience…", "dividing between relevant and irrelevant content. It is necessary to develop skills of conveying messages"*, and workshops for strengthening the student's self-confidence: *"work on self-confidence, body language…", "imparting tools that will help acquire self-confidence and deliver the presentation optimally",*

About 26% of the respondents aimed their criticism/suggestions not necessarily at the system but rather inwards, emphasizing the steps that they themselves should take to improve their personal capacity to give a presentation. For instance, practicing before the presentation: *"Preparing and practicing face-to-face before giving the presentation on Zoom in the course"*, preparing before writing the slideshow: *"to first read lots of information"*, good planning and emphases: *"write down what you want to say, prepare a good slideshow, rehearse before giving the presentation, and prepare everything in advance and not at the last moment In addition, preparing a slideshow that will connect the audience to the topic, pictures and less text in the slideshow".*

Students also noted the need for direction and guidance to improve technical skills of using aids and tools for preparing a slideshow, as well as general knowledge of preparation, planning, and presentation: *"showing how to build a good presentation on a university level and more thorough guidance on the topic", "better learning of the capabilities of new media tools", , "improving the ability to present, plan time, and choose the presentation method".* A small proportion of the respondents suggested adaptations of the presentation environment as a tool for improving capabilities of presenting slideshows, namely, reducing the number of viewers while presenting the slideshow: *"that the presentation will be only to the lecturer", "small groups of students to whom the presentation is given, so that it will be less stressful"*. A small proportion noted that presentations should be avoided altogether: *"not give presentations, but rather exams", "cancel the presentations" (*Table 10).

Table 10: Analysis of themes - fifth question (n=64)

Themes	%	n
Focused instructions and guidance by the lecturer	37.9	25
Self-criticism (ways of self-improvement)	25.8	17
Workshops on rules	24.2	16
Workshops to improve technological capabilities	13.6	9
Cancelling presentations	7.6	5
Adapting the presentation environment	4.5	3
Giving examples of presentations	4.5	3
Evaluation of presentation/ preparation time	4.5	3

It is evident from the analysis that the participants contend that clear guidance by the lecturers, imparting relevant tools and skills, encouraging awareness and personal responsibility among students, improving the technical aspects, and adapting the presentation environment, can contribute significantly to improving the process of preparing and presenting slideshows and to increasing their success with the assignment.

In the sixth question, the students were asked to express their satisfaction with the use of presentations as an assessment tool: "Are you satisfied with the use of presentations as an assessment tool in the course?". Sixty-five of the 66 participants answered the question (98.5%). More than half the respondents (approximately 59%) answered that they are satisfied. About 8% noted that the presentation helps reach deeper learning of the material: *"Yes, it makes it possible to investigate the topic presented in depth…" "…You learn a million times more from an assignment than from a course…"* and express their proficiency in the material: *"Yes, an exam does not check the material, an assignment does", "Yes, it allowed me to present what I understand and to convey it to the lecturer and students"*. Students who expressed satisfaction also noted that the presentation contributed to developing creative thinking and interest: *"Yes, it allows you to investigate the topic presented in depth and to share it with others in an interesting way",* as well as the ability to overcome difficulties: *"Yes, I think that it's a tool that teaches how to confront a stressful situation".*

Approximately 27% of the respondents expressed dissatisfaction with the use of presentations as an assessment tool; 12% reported that the assignment had caused them emotional difficulties such as pressure: *"No, in my opinion presentations cause unnecessary stress and don't really help you understand the material"* and harm to their self-confidence: *"Sort of. On the one hand you learn a million times more from an assignment than from a course. On the other, for those who have social anxiety it reduces their confidence"*. A small proportion of the students who expressed dissatisfaction complained that the presentation is unsatisfactory due to the imbalance between its contribution to their achievements in the course and the efforts put into preparing and presenting the slideshow: *"…and not when in practice it adds a point or two to the final grade although much work is put into it"* as well as its contribution to understanding the study material: *"No, in my opinion presentations cause unnecessary stress and don't really help you understand the material".*

Approximately 12% of the students expressed partial satisfaction with presentations as an assessment tool. Some contended that the assignment should be combined with a final paper *"It depends, if it's combined with a final paper…"* or offered as an elective, as an alternative assessment tool: *"It's nice as an elective. Some people find it hard to speak to an audience and would prefer not to give a presentation…"* (Table 11).

Table 11: Analysis of themes - sixth question (n=65)

Themes	%	n
Yes	59.1	39
No	28.8	19
Partially	12.1	8
Improving capabilities (creativity and handling difficulties)	9.1	6

Themes	%	n
Improves proficiency in the material	7.6	5
Allow self-expression	6.1	4
Allows conveying a message	6.1	4
Emotional difficulties	4.5	3
Unsatisfactory	4.5	3

From the analysis shown it appears that most of the students expressed satisfaction with use of presentations as an assessment tool in the course. The students noted several benefits of using presentations, including more thorough learning of the material, the possibility of showing proficiency in the study material, developing creative thinking and interest, and the ability to overcome difficulties such as facing an audience.

At the same time, slightly less than one third of the respondents expressed dissatisfaction with the use of presentations. The reasons included causing emotional difficulties such as stress and harm to one's self-confidence, the feeling that the assignment does not contribute enough to the final grade considering the necessary efforts and also doubts regarding its contribution to understanding the study material. A relatively small proportion expressed partial satisfaction and suggested combining the presentation with other assignments or offering it as one of several choices. In summary, although most of the students saw presentations as a beneficial assessment tool that contributes to learning and to personal development, some pointed to their challenges and shortcomings.

In the seventh question, the participants were asked to state their opinion about the possibility of utilizing the presentation assignment as an exclusive assessment tool: "Would you prefer the presentation to be the exclusive assessment tool in the course? Explain…". All participants but one answered the question. The aggregated answers indicate that a large majority of the respondents, about 65%, were not in favor. Of these, some 32% claimed that the presentation assignment does not sufficiently reflect all aspects necessary for evaluating the student, as evident from the direct quotations: *"No, because giving a presentation still does not say anything clear about the student".* Others added that presentations as an assessment tool do not reflect their capabilities: *"No. Because I'm sure that there are many other students like me who find it less comfortable to speak to an audience or who cannot do it at all, so I think that presentations should not be an exclusive assessment tool in the course", "No, because sometimes there are students who find presentations difficult and who do not have the ability to prepare a presentation that meets the lecturer's requirements".*

Approximately 18% of the respondents expressed a preference for combining presentations as assessment tools with other tools such as writing a paper or an exam: *"No, because I think that there is also need for a more theoretical detailed paper that is the basis for the shortened presentation".*

A similar proportion noted that presentations as an assessment tool might discriminate against students who find it hard to stand before an audience. *"No. Because I'm sure that there are many other students like me who find it less comfortable to speak to an audience or who cannot do it at all, so I think that presentations should not be an exclusive assessment tool in the course".*

Another claim that arose from the participants' answers is that presentations are not suitable as exclusive assessment tools for various reasons, such as lack of comprehensibility, *"No, it doesn't allow you to display everything"*, lack of contribution to understanding the material: *"No, in my opinion it is a terrible assessment tool, the course material is not truly internalized"*, limited time restrictions: *"No, because a presentation is 10 minutes at most and it is not possible to assess an entire semester based only on a presentation"*, and lack of clarity regarding assessment criteria: *"I need to know more details to answer that question. What exactly would they assess…"*

Other students expressed reserved consent with the exclusive use of presentations as an assessment tool, depending on the type of course: *, "It depends on the course and the number of students"* and of the presentation duration: *"Yes, if there is enough time", "I would be very glad, but I think that there is a problem with the time, because when I give presentations the time is usually limited to 10 minutes and that is not really sufficient to convey the knowledge acquired".* Only about 17% of all respondents expressed consent with using presentations as an exclusive assessment tool, in the claim that it helps and improves the learning experience through experiencing. *"Yes… in my opinion it's the best way for students to show what they are learning and also to* *convey their views", "Yes, it imparts experience with an audience and learning the study material by processing and not by revision"* (Table 12).

Table 12: Analysis of themes - seventh question (n=65)

Themes	%	n
Does not reflect all aspects	31.8	21
Does not reflect capabilities	21.2	14
Preference for combining with other assessment tools	18.2	12
Discriminates due to concerns of facing an audience	18.2	12
Is inappropriate as an assessment tool (time and comprehensiveness)	16.7	11
Does not contribute to understanding the material	13.6	9
Depends on the time and place	10.6	7
Splitting into small assignments throughout the course	4.5	3
Yes, constitutes a tool for learning based on experiencing	16.7	11

The analysis presented shows that most of the students objected to using presentations as an exclusive assessment tool in the course. Their main claim is that presentations do not reflect all the student's capabilities and knowledge and might discriminate against those who find it hard to face an audience. Some of the students noted additional shortcomings of presentations as an assessment tool, such as the time restriction, lack of clarity regarding assessment criteria, and lack of contribution to understanding the study material.

Nevertheless, several students suggested different ways of improving presentations as an assessment tool, for instance by combining them with other tools such as papers or exams, adapting them to the specific course type, lengthening the presentation time, and splitting the assignment into several short presentations throughout the semester. Only a relatively small minority of the students supported exclusive use of presentations as an assessment tool, particularly due to their contribution to improving presentation skills and experiential learning.

In conclusion, the main insight is that despite the potential benefits of presentations, most students contend that they cannot be relied on as a single and inclusive means of assessment. It is also evident that, according to the respondents, in order to integrate them more significantly in the assessment process it is necessary to think of adaptations and improvements that will give proper expression to students' different capabilities, while reducing its shortcomings as an assessment tool.

Summary and conclusions from analysis of the responses to the open-ended questions

The overall analysis of the open-ended questions indicates participants' contention that use of presentations as an assessment tool in academic courses has significant benefits alongside challenges and shortcomings. Most of the students noted that presentations contribute to enhancing understanding of the study material, allow self-expression and creativity, and help develop skills such as facing an audience.

Nonetheless, a considerable portion of the students raised challenges stemming from concerns and difficulties involving facing an audience, trouble condensing the material and adapting it to a slideshow format, and lack of clarity regarding instructions and assessment. Accordingly, a large majority of the students objected to use of presentations as an exclusive assessment tool, claiming that it does not truly reflect students' full range of capabilities and might discriminate against students who find it hard to face an audience.

The students suggested several ways of improving presentations as an assessment tool, such as giving clear instructions and closer guidance, holding workshops for imparting relevant skills, combining presentations with other assessment tools, and adapting them to the nature of the course. Hence, it is evident from the findings that there is room for developing and improving presentations as an assessment tool, while providing a response to needs and challenges raised by the students and gradual assimilation of advanced technology.

4. Summary and Conclusion

Overall analysis of the quantitative and qualitative (open-ended questions) research findings indicates a correspondence between the findings in the two types of analysis. The quantitative analysis points to a positive association between high self-evaluation by students and a sense of comfort with face-to-face presentations and the perception of presentations as improving learning capabilities. Similarly, in the qualitative analysis most of the students expressed satisfaction with use of presentations as an assessment tool and noted benefits such as enhancing understanding, improving retention, and contribution to more thorough and significant learning.

Furthermore, the quantitative analysis indicates the importance of attendance in classes where presentations are given for succeeding in one's studies; in the qualitative analysis as well the students noted the contribution of interactions during the presentation to the learning process. In addition, the hierarchical regression in the quantitative study emphasizes the importance of integrating various assessment tools alongside the presentations, while in the qualitative findings as well the students objected to presentations as an exclusive assessment tool and suggested that they be combined with other tools.

Other general insights arising from the findings:

- Presentations have significant potential to contribute to enhancing learning beside developing important skills, but they cannot be an exclusive assessment tool.
- Factors such as self-esteem, self-efficacy, as well as attendance and interaction during presentations, affect the effectiveness of the presentation.
- It is necessary to think about adaptations in order to transform presentations into more effective assessment tools, such as clear instructions, imparting tools, improving technical aspects, and adaptation to students' needs.

In conclusion, the integrated findings stress the complexity of presentations as tools for learning and assessment and the need for thorough consideration of how to realize their full potential while addressing their limitations. In current processes of teaching, learning, and assessment, the research findings shed light on presentations using slideshows as tools and means that support learning. The research findings may contribute to improving students' learning processes and evaluating their academic and personal development. For this purpose, however, it is necessary to train students to experience the preparation of presentations by effective slideshows and to face an audience, and it is also crucial to develop a culture with a cooperative-learning climate appropriate for an era of digital innovation.

Delivering presentations and facing an audience are essential skills in the modern workplace. Students will find themselves in professional situations that require presentations, usually using slideshows and conveying messages to an audience of listeners or viewers. Facing an audience and presentation skills can be learned and there are even ways of successfully dealing with fear and reinforcing confidence in such situations. Good techniques will grant the necessary tools for delivering presentations in a more efficient and focused manner and increasing the chances of attaining the goals of the presentation.

Ethics Statement: Ethical approval for this study was obtained from Ariel University of Samaria Institutional Review Board (Approval number AU-SOC-ND-20240814).

AI Statement: The authors declare that they have not used any type of generative artificial intelligence for the writing of this manuscript, nor for the creation of images, graphics, tables, or their corresponding captions.

References

Bar, Y. (2010). "Everything is predetermined by there is freedom of choice: *Al Hagova, 5*, 46-48. [Hebrew]

Carmine, G. (2018). T*alk like TED: The 9 public-speaking secrets of the world's top minds*. St. Martin's Press.

Fisk, G. D. (2019). *Slides for students: The effective use of PowerPoint in education*. University of North Georgia Press.

Gallo, C. (2014). *The presentation secrets of Steve Jobs: How to be insanely great in front of any audience.* McGraw Hill.

Grieve, R., Woodley, J., Hunt, S. E., & McKay, A. (2021). Student fears of oral presentations and public speaking in higher education: A qualitative survey. *Journal of Further and Higher Education, 45*(9), 1281–1293. https://doi.org/10.1080/0309877X.2021.1948509

Hadiyanti, K. W., & Widya, W. (2018). Analyzing the values and effects of powerpoint presentations. *LLT Journal: A Journal on Language and Language Teaching*, *21*(Suppl), 87-95.

Hammer, R. & Ankori, G. (2008). Designing Powerpoint slideshows as navigation and concept maps. *Al Hagova, 7*, 38-40. [Hebrew]

Jacobson, M., & Levin, J. (1995). Conceptual Frameworks for network learning environments and hypertext: Frameworks for constructing personal and shared knowledge spaces. *International Journal of Educational Telecommunications*, *1*(4), 367-388.

Kozminsky, E., Gad, A., & Barkai, I. (2000, February 13-14). E*valuating slideshows produced by students: What do students, teachers, supervisors, and novice teachers say about evaluating slideshows and what do they mean?* [Presentation]. Opening Gates in Teacher Training Conference, Mofet Institute, Israel. [Hebrew]

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. Educational *Psychology Review*, *19*, 309-326. http://dx.doi.org/10.1007/s10648-007-9047-2

Margaret, D. (2015). Powering pedagogy. *Language Magazine*, *14*(11), 20–23. https://www.languagemagazine.com/powering-pedagogy/

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, *6*(4), 359-377.

Meishar-Tal, H. (2011). Development and operation of a cooperative study assignment in online distance learning. In D. Chen & G. Kurtz (Eds.), *ICT, Learning and Teaching* (pp. 451-467). The College for Academic Studies. [Hebrew]

Osman, N., Noor, S. S. M., Rouyan, N. M., & Hat, N. C. (2022). The use of PowerPoint in developing multimedia-based teaching and learning materials for learning Arabic language. *Journal of Language and Linguistic Studies, 18*(1), 15-21. doi:10.52462/jlls.163

Lee, K., Kang, E., & Park, E. J. (2023). Storytelling as a learning tool in creative education: A case study in an architecture design studio. *Thinking Skills and Creativity*, *48*, 101274.

Salant, A. (2022). What happens when students prepare their own presentations? *Digital Knowledge and Curiosity*. http://www.amisalant.com/?p=24009 [Hebrew]

Tannenbaum, S. I., Smith-Jentsch, K. A., & Behson, S. J. (1998). Training team leaders to facilitate team learning and performance. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 247–270). American Psychological Association. https://doi.org/10.1037/10278-009

Widhayanti, A., & Abduh, M. (2021). Penggunaan Media Audiovisual Berbantu Power Point Untuk Meningkatkan Hasil Belajar Peserta Didik di Sekolah Dasar. *Jurnal Basicedu*, *5*(3), 1652-1657.

Wecker, C. (2013). interview. https://www.lmu.de/en/newsroom/index.html

How Social Media Marketing Drives e-Learning Platform Adoption: A Multigroup Approach

Inés Requez-Cipriano and Martín Mauricio-Andía
Department of Communications, Universidad Peruana de Ciencias Aplicadas, Lima, Perú

u201113556@upc.edu.pe
pccmmmau@upc.edu.pe (corresponding author)

https://doi.org/10.34190/ejel.23.2.3595

Abstract: Technological advancements have been instrumental in the evolution of the educational landscape, particularly with the proliferation of electronic learning platforms, which have witnessed significant growth in recent years, offering individuals a diverse array of customised courses. In this vein, the present study seeks to elucidate the impact of social media marketing activities on electronic word-of-mouth and student intention to engage with these platforms. Specifically, we aim to investigate how activities such as entertainment, interaction, trend analysis, personalisation, and advertising, conducted by leading digital e-learning platforms on social networks, shape user behaviour. Furthermore, we endeavour to assess the mediating role of electronic word-of-mouth, with a particular emphasis on the educational attainment level of respondents. To achieve these objectives, we conducted a cross-sectional quantitative study involving a sample of 303 participants, all of whom were users of at least one electronic learning platform. Structural Equation Modelling (PLS-SEM) was employed to analyse the proposed model comprehensively. Our findings underscore the substantial influence of social media activities on both electronic word-of-mouth and students' propensity to utilise these platforms. Notably, our multi-group analysis, stratified by educational level, reveals nuanced patterns. Specifically, university and technical users demonstrate a heightened association between social media marketing activities, electronic word-of-mouth, and intention to use, reflecting their urgency in accessing the service and thereby exhibiting a greater receptivity to digital platform activities. Conversely, postgraduate users place greater significance on electronic word-of-mouth, underscoring the importance of authentic user perceptions in shaping their engagement decisions.

Keywords: e-Learning, Electronic word of mouth, Intention to use, Social media marketing activities, Multigroup analysis

1. Introduction

The accelerated development of information and communication technologies, the Internet, and web-based applications has given rise to e-learning (Liu, Liao and Pratt, 2009). E-learning encompasses various forms of teaching and has become an effective means of facilitating online learning processes (Troussas, Krouska and Sgouropoulou, 2021). It has gained great acceptance due to benefits such as cost-effectiveness, collaboration, personalized learning, and others and has become the learning system for future generations (Talebian, Mohammadi and Rezvanfar, 2014).

The growth of e-learning in recent years in global markets has improved corporate competencies and individual talent preparation, with government initiatives promoting the entry of global companies into the market (Ghewari and Anute, 2021). The global e-learning market is projected to reach over $243 billion in the next few years (Korhonen, 2024), with the United States and Europe currently leading the way in this area (Barbour *et al.*, 2020). However, the impact of e-learning is not limited to these regions; in Latin America, the number of users of learning platforms has grown exponentially, reaching figures in the thousands in some countries. This trend highlights the relevance of studying users' motivations and their interaction with learning platforms (Stanley and Montero Fortunato, 2022). Therefore, understanding students' perceptions is essential for identifying key opportunities for improvement in the relationship between users and e-learning platforms, which in turn can facilitate the adoption of effective strategies for student retention and the enhancement of the online learning experience (Mukhtar *et al.*, 2020; Sun *et al.,* 2008)

One of the key factors in assessing the perceived acceptance of e-learning is social media marketing activities, facilitated by the sense of community that social networks provide academically and professionally (Al-Shdayfat, 2018). This type of marketing has been identified as an important catalyst for the development of user loyalty towards educational platforms (Mujica, Villanueva and Lodeiros-Zubiria, 2021). In social media, electronic word-of-mouth (WOM) is generated as a result of users' interactions with educational brands, subsequently affecting their intention to use. Users' intentions are influenced by the opinions of others, and there is the possibility of receiving a positive response from the company if a favourable perception of the brand is obtained (Erkan and

Cite this article: Requez-Cipriano, I. and Mauricio-Andía, M. 2025. "How Social Media Marketing Drives e-Learning Platform Adoption: A Multigroup Approach", *Electronic Journal of e-Learning*, 23(2), pp 46-63, https://doi.org/10.34190/ejel.23.2.3595

Evans, 2016; Goh *et al.*, 2017). This phenomenon is particularly relevant in the context of e-learning, where trust in learning platforms can determine technological adoption (Zandi, Lahrash and Shakhim, 2022).

Over the years, research on the relation between social media marketing activities and the intention to use has been developed in various industries (Kim and Ko, 2012; Chen and Lin, 2019). Other studies have highlighted the importance of social media marketing activities in generating electronic WOM (Sharma *et al.*, 2021). From a theoretical perspective, this relationship can be explained through the Uses and Gratifications Theory (U&G) (Katz, Blumler and Gurevitch, 1973), which posits that individuals use media based on their needs and expectations, obtaining specific gratifications. In the context of e-learning, marketing activities on social media can influence usage intention by fulfilling users' informational, social, and entertainment needs. By providing valuable content, facilitating interaction with other users, and creating engaging learning experiences, these activities can strengthen participation and commitment to digital platforms. This theoretical approach further justifies the examination of digital marketing activities as a key element in the e-learning landscape.

In the field of education, emphasis has been placed on the effectiveness of social media marketing activities in accepting online education at different levels of education (elementary, technical, and university) (Shams *et al.*, 2022; Shehzadi *et al.*, 2021). In addition, its importance in the digital ecosystem is underscored by its connection with the intention to use, reducing the user abandonment rate (Tan *et al.*, 2014). Similarly, the exposure of social media marketing activities to electronic WOM is noted, underscoring the importance of its relationship with learners' genuine perceptions of learning platforms (Shehzadi *et al.*, 2021).

However, studies on the mediating role of electronic WOM have been neglected. Some references have mentioned its positive influence as an intermediary among satisfaction, loyalty, and fidelity (Hasan, Al-Dmour and Al-Dmour, 2020; Wei *et al.*, 2021). This emphasizes its importance in the exchange of online options generated by social media marketing activities and its impact on the intention to use in the academic context (Al-Shdayfat, 2018; Sharma *et al.*, 2021).

Despite these findings, the role of eWOM in mediating the relationship between social media marketing activities and usage intention in e-learning platforms has not been sufficiently explored, representing a gap in the literature that this study seeks to address. In this context, the present study has two main objectives: (1) to analyse the relationship between social media marketing activities (SMMA), electronic word-of-mouth (eWOM), and usage intention (UI) in e-learning platforms, and (2) to evaluate the mediating role of eWOM in this relationship. The justification for this analysis lies in the recognition that, in the highly competitive and rapidly changing e-learning environment, platforms must explore the underlying motivations of users to optimize their adoption. To achieve these objectives, the following research questions will be formulated:

- RQ1: To what extent do social media marketing activities influence the usage intention of e-learning platforms?
- RQ2: What is the impact of eWOM on the usage intention of these platforms?
- RQ3: Does eWOM act as a mediator in the relationship between social media marketing activities and usage intention in the context of e-learning?

This research will contribute to the field of e-learning by providing empirical evidence on the impact of digital marketing strategies on the acceptance of these platforms, enabling educational institutions and online service providers to optimize their student acquisition and retention strategies.

2. Literature Review and Hypotheses Development

2.1 Social Media Marketing Activities

Social media marketing is a space that involves information and ideas through social networks, promoting the feeling of closeness (Ibrahim, Aljarah and Sawaftah, 2021). It allows for the sharing of common interests, thoughts, and ideas among people (Li, Larimo and Leonidou, 2021). This leads consumers to follow the brand through online media (Chen and Li, 2019) enabling them to connect with organizations and stimulate their engagement through the creation of communities in real-time or asynchronously (Carr and Hayes, 2015). The goal is to create value and achieve a significant impact on their reputation (Kim and Ko, 2010).

In recent years, research has been conducted to analyse the acceptance of social media marketing activities and their impact on users' responses. This involves their involvement in the analysis process to define the intention to use (Wibowo *et al.*, 2021), demonstrating its contribution to creating this intention among consumers and communities in social networks (Cheung, Pires and Rosenberger, 2020). This impact has been studied across various fields and environments such as social networks (Chen and Lin, 2019), e-commerce (Yadav and Rahman, 2017). These studies have revealed direct effects on business results, sales volume, profits, and growth rates (Bronnenberg, Kim and Mela, 2016). This is attributed to the constant communication it facilitates with consumers (Breitsohl, Kunz and Dowell, 2015), collectively enabling the effective execution of social media marketing activities as an integral part of a marketing strategy (Li, Larimo and Leonidou, 2021). This enables companies to acquire a functional role to frame, define, and effectively implement activities on these platforms (Kim and Ko, 2012).

In the field of e-learning, a sense of community in social networks is key to its success (Khurshid *et al.,* 2023). Research has mentioned that students exhibit intentions to use influenced by social networks in terms of academic and professional contexts (Al-Shdayfat, 2018). This influence is attributed to the valuable information provided by social media marketing activities available on the social networks of e-learning platforms (Ghewari and Anute, 2021). In addition, interaction with social media marketing activities in the context of e-learning can significantly improve student participation and enable active feedback (Troussas, Krouska and Sgouropoulou, 2021), fostering alliances in e-learning where people create and maintain personal and commercial relationships (Vate-U-Lan and Masouras, 2018).

In the business context, social media marketing activities allow for a more comprehensive understanding of significant digital trends that e-learning platforms should incorporate into their marketing strategy to stand out (Kumar, 2021).

One of the initial studies on social media marketing activities has grouped them into entertainment, interaction, trending, personalization, and WOM communication within the premium brand category (Kim and Ko, 2012). Subsequently, the same activities were expanded to include entertainment and promotion in a study focusing on their implementation in a brand's social networks in the consumer category (Bilgin, 2018). In 2017, other authors incorporated this information in a study that tested its influence on brand equity and the intention of use in e-commerce (Yadav and Rahman, 2017). Later, activities such as entertainment, interaction, trend, personalization, and perceived risk were grouped in a study on the airline industry, revealing that people's engagement and intention to use can increase with frequent utilization (Seo and Park, 2018).

As can be observed, research has been conducted on the acceptance of social media marketing activities in different settings (Chen and Lin, 2019).

Therefore, for this study, the aforementioned research is taken as a reference. The dimensions of social media marketing activities are classified as follows: entertainment, interaction, advertising, personalization, and trend.

Entertainment

Entertainment in social networks has a great power to satisfy needs, as it allows a strong emotional connection between the brand and consumers (Panigyrakis, Panopoulos and Koronaki, 2020). It is one of the most important activities in social networks because it facilitates acceptance within the integration of a group of people (Khan *et al.*, 2022). This could create relationship with a brand, generating a positive WOM and possible intention to use it (Ibrahim, Aljarah and Ababneh, 2020; Seo and Park, 2018). In addition, users could positively generate electronic WOM as a declaration of their endorsement (Liu, Shin and Burns, 2021).

In the context of e-learning, entertainment enables a better adoption of e-learning (Yan, Eng and Seong, 2024) due to the state of attention it generates in people, similar to the one generated by online games (Lashari *et al.,* 2024). Likewise, entertainment is an indicator of learning performance because the use of social networks is ideal for knowledge sharing and student learning (Eid and Al-Jabri, 2016).

Interaction

Several authors have agreed that interaction within social networks motivates users to create their content and share it online (Kamboj *et al.,* 2018; Hajli *et al.,* 2017). Social networks must be perceived as sincere to consistently convey the brand's values and generate electronic WOM (Hennig-Thurau *et al.*, 2004). This is achieved by fostering customers' understanding through their interaction with the brand (Cheung, Pires and Rosenberger, 2020), enabling the entry of new customers through the exchange of opinions (Lee and Shin, 2014). Accordingly, social networks are designed to persuade online social interaction (Xu, Lovett and Law 2022). Positive emotions generated in the user increase willingness to engage in interactions, thereby triggering intentions to use the brand (Kramer, Guillory and Hancock, 2014). In the context of e-learning, interactions on social networks allow learners to easily exchange ideas, documents, and audio-visual material, forming an effective participatory community (Greenhow and Lewin, 2016; Salloum and Shaalan, 2019).

Advertising

Advertising is a media activity performed by a company or agent in one or more channels (Stephen and Galak, 2012) to persuade a close environment to make a purchase (Stafford et al., 2022). Social networks are those that establish a relationship between professionals and other consumers to create new opportunities to increase brand and consumer awareness through advertising (Tran, Muldrow and Ho, 2021). Subsequently, it became the main generator of positive perceptions that simultaneously increase the brand image value and intention to use (Rusfian and Alessandro, 2021) . The ability of advertising to generate electronic WOM on different platforms helps to effectively position the brand image (Kaplan and Haenlein, 2010). In the context of e-learning, advertising is part of the effective marketing products that help educational institutions attract and engage students (Khan and Joshi, 2006). Given its influence on building user attitude, it can be used to provide information to potential and former users about the benefits of using e-learning (Lee, 2010).

Personalization

Personalization is the extent to which services are tailored to meet consumers' personal preferences (Godey *et al.*, 2016). Several authors have mentioned that personalization in social networks is a tool that should provide attractive and useful information (Kim and Ko, 2010; Merrilees, 2016). Brands may personalize their online communication for certain people who frequently interact on the brand's social networks (Yadav and Rahman, 2018). Through personalization in social networks, users can freely express their thoughts and generate electronic WOM for the brand (Kim and Ko, 2010). The importance of flexibility in personalizing some products or services is considered to show an understanding of the intention to use by customers toward the brand and to garner a positive response (Anshari *et al.*, 2019). In the context of e-learning, it is convenient to develop personalized services to endorse users' intention to use (Kang, 2024), considering the positive impact of personalization in consumers' engagement with the brand (Shanahan, Tran and Taylor, 2019).

Trend

A brand's trend is defined as the speed with which it communicates current information (Naaman, Becker and Gravano, 2011). This capability captures customers' attention and associates the brand with positive feelings, thereby strengthening customer loyalty (Liu, Shin and Burns, 2021; Panigyrakis, Panopoulos and Koronaki, 2020). Furthermore, it contributes to the construction of a positive brand experience, enabling the perception of leadership. This, in turn, motivates others to learn about the brand and generates electronic WOM through discussions and comments (Chan *et al.*, 2014).

Studies have revealed that interaction through electronic word-of-mouth (eWOM) and digital trends directly influence consumer engagement with the brand, thereby strengthening brand recognition and generating intention to use (Cheung, Pires, and Rosenberger, 2020). However, while these findings are relevant, the existing literature in the context of online learning platforms (e-learning) has not adequately explored the specific factors influencing user adoption and engagement within this environment.

In the context of online learning, staying abreast of the latest technological trends is crucial, as users value innovative e-learning offerings, as highlighted by the adoption of emerging technologies that enhance the user experience (Cook and Triola, 2014). This involves integrating future digital tools into the learning process, such as augmented reality, which has shown a positive impact on student engagement and retention (Düking, Holmberg, and Sperlich, 2018).

Furthermore, the literature suggests that eWOM in the context of e-learning has the potential to modify students' perceptions of the quality and effectiveness of learning platforms. Experiences shared by other users, particularly on social platforms, can generate positive expectations that drive intention to use and student satisfaction (Wang *et al.,* 2023)**.** This dimension of eWOM requires more thorough analysis in the context of online learning, given its potential impact on the choice of educational platforms.

In conclusion, while existing research provides a general understanding of the influence of eWOM and technological trends on consumer engagement, there is a need for a more focused and critical analysis of the contextual factors affecting student adoption, engagement, and satisfaction on e-learning platforms.

Accordingly, the following hypotheses are postulated:

H1: Social media marketing activities impact the formation of positive WOM in e-learning users.

H2: Social media marketing activities positively impact e-learning users' intention to use.

2.2 Electronic Word-of-Mouth

WOM is defined as an informal communication about the use or characteristics of products or services shared with other customers (de Matos and Rossi, 2008). It influences the development of consumer attitudes and their intentions to use products or services (Chatterjee and Kumar Kar, 2020). With the advent of the Internet, electronic WOM has been referred to as the communication by potential or actual customers of a brand or company that maintains a presence in social networks (Chu, Lien and Cao, 2019; Kara *et al.*, 2018). This communication fosters closeness, loyalty, and an emotional relationship between brands and consumers (Brodie *et al.*, 2013). Consumers rely on this communication before making a product purchase. They seek information gathered online (Alrwashdeh, Emeagwali and Aljuhmani 2019), leading to the generation of conversations where they express their perceptions about the quality of the product and/or service through electronic WOM. This information is valuable in predicting their intentions to use (Jalilvand and Samiei, 2012; Zhao *et al.*, 2016).

The mediating role of electronic WOM is manifested as an important element between brand image and intention to use in consumer markets (Jalilvand and Samiei, 2012). It influences customers' value perception and loyalty intentions due to the exchange of opinions among customers (Kim and Hyun 2019). Studies have revealed that the generation of this communication positively impacts as an intermediary between satisfaction and loyalty of the e-consumer (Hasan, Al-Dmour and Al-Dmour, 2020).

However, in the e-learning context*,* only few studies have focused on the mediating analysis of electronic WOM. Nevertheless, the studies found in this context have revealed the important role of electronic WOM, as educational brands are more exposed to their users generating their opinions (Shehzadi *et al.*, 2021). Furthermore, e-learning presents a significant relationship with electronic WOM, as it allows the brand to have a positive perception by its students (Cole, Shelley and Swartz, 2014). Therefore, this study seeks to analyse the mediating role of electronic WOM between social media marketing activities and intention to use to know its level of influence between both variables in the context of e-learning.

Accordingly, the following hypothesis is postulated:

> *H3: Electronic WOM presents a mediating role between social media marketing activities and e-learning users' intention to use.*

2.3 Intention to use

In the marketing field, this variable has been extensively researched as a determinant of actual technology use in the context of technology adoption (Hoffman *et al.*, 2022; Lin *et al.*, 2020). Within the context of social networks, research has revealed that the content generated by social media marketing activities motivates interaction (Aji, Nadhila and Sanny, 2020), enabling seamless two-way communication for sharing information and opinions. This openness allows brands to be more transparent with consumers compared with traditional media, thereby increasing brand engagement and strengthening consumers' intention to use the brand (Yadav and Rahman, 2017).

In the context of online learning, the likelihood of adopting a learning platform increases when individuals perceive positive feedback through electronic word-of-mouth (e-WOM) (Noh, Jang and Jeon 2021). This suggests that user experiences, as communicated through e-WOM, play a significant role in shaping the intention to engage with these platforms (Zhang *et al.,* 2022).

While there is a substantial body of research examining the general relationship between e-WOM and consumer behaviour across various industries, the specific dynamics within the online learning sector remain underexplored. This relationship has been studied in diverse contexts, including the consumer industry (Aji, Nadhila, and Sanny, 2020), e-commerce (Goh et al., 2017), and social media (Erkan and Evans, 2016). However, the educational context, particularly online learning platforms, presents unique challenges and opportunities. Research indicates that online reviews are pivotal in influencing the adoption of online learning platforms and enhancing student satisfaction (Rabah *et al.,* 2024).

A more focused analysis is required to investigate the specific challenges and opportunities within online learning environments. Key factors, such as digital literacy, user interface design, and the role of social influence through e-WOM, warrant a more critical examination. Moreover, understanding how different types of e-WOM (e.g., peer reviews, instructor feedback) affect student engagement and their ongoing use of learning platforms could offer valuable insights for improving platform design and increasing student participation. A deeper exploration of these contexts will not only enrich the current literature but also contribute to the development of more effective strategies aimed at enhancing student interaction with online learning platforms.

Accordingly, the following hypothesis is postulated:

H4: Electronic WOM positively impacts e-learning users' intention to use.

Figure 1 below illustrates the hypotheses proposed in this research. The diagram serves as a visual representation of the hypotheses analysed in the study.

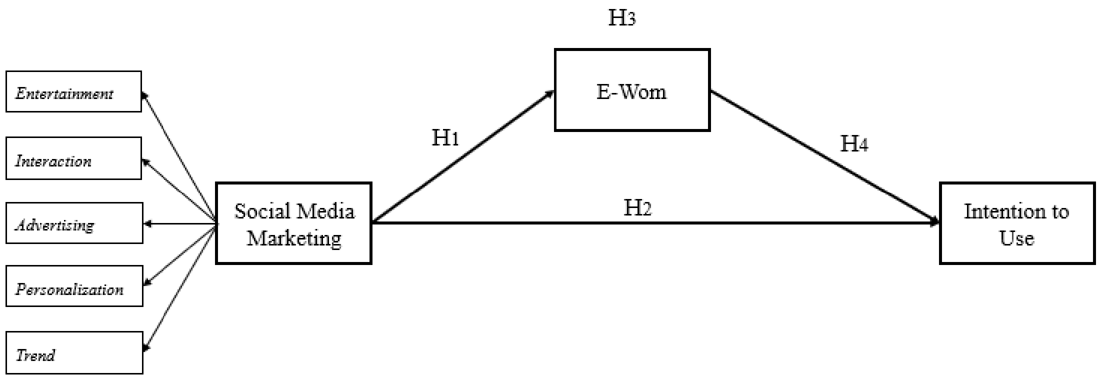

Figure 1*: Research model*

3. Method

3.1 Data Collection

The study population comprised technical, university, and postgraduate students who had taken an online course on an e-learning platform in the last six months. This last criterion was essential for completing the questionnaire; those who did not comply with this requirement were directed to complete it. The sampling procedure adopted was non-probabilistic, specifically convenience sampling, due to time and resource constraints in data collection. This approach allowed for quick access to a diverse group of participants based on their recent experience with online learning platforms (Etikan, 2016).

The sample comprised 303 people who had taken at least one e-learning course through an online platform(s) in the last six months. The most commonly used platforms included Udemy, Domestika, Coursera, and Crehana, selected for their popularity in technical, university, and postgraduate contexts. The demographic profile of the sample encompassed women (51.49%) and men (48.51%), with ages ranging from 18 to 25 years old (41.58%), and a significant portion studying and working (28.38%). Similarly, the educational levels of the participants varied, with the majority having a university background (55.00%), followed by technical (34.33%), and postgraduate (10.67%) qualifications. Table 1 presents the demographic results of the study.

Table 1*: Demographic results*

Variable	Category	Frequency	Percentage
Genre	Men	147	48.51%
	Women	156	51.49%
Age range	Under 18 years old	16	5.28%
	18 to 25 years old	126	41.58%
	26 to 35 years old	123	40.59%
	36 to 45 years old	32	10.56%
	46 to more	6	1.98%
Educational level	Technical	104	34,33%
	University	167	55,00%
	Post Graduate	32	10,67%

Variable	Category	Frequency	Percentage
Occupation	Studying	25	8,25%
	Working	85	28,05%
	Studying and working	86	28,38%
	Leading my own business	64	21,12%
	Looking for a job	43	14,19%

3.2 Measures

A questionnaire of 18 questions adapted to the field of study was translated into Spanish. For each question, a Likert-type scale format was used, comprising five (5) levels ranging from "strongly disagree" (1), "disagree" (2), "neither agree nor disagree" (3), "agree" (4), and "strongly agree" (5). The questionnaire was translated following a back-translation approach to ensure semantic and conceptual equivalence between the original and translated versions (Beaton *et al.,* 2000).

In the case of social media marketing activities, five dimensions were included: entertainment, interaction, trend, advertising, and personalization, making two items for each dimension, adapted from the study that aimed to assess the perception of each of them in social networks (Kim and Ko, 2012). For electronic WOM, four indicators were used, referring to a study with a similar objective, to assess its role in the electronic context (Hidayanto *et al.*, 2017). Likewise, for intention to use, four indicators were used, adapted from the study that sought to understand the impact of user behaviour on their intention to use online platforms (Alraimi, Zo and Ciganek, 2015). The use of the Internet for the research study through surveys was chosen due to the advantages it offers, enabling quick and cost-effective outreach to a larger audience, facilitating the adaptation of questions, and simplifying the management of data analysis (Evans and Mathur, 2005). However, this approach has limitations related to self-selection bias, as participants must have access to Internet-connected devices, which could exclude certain segments of the population.

4. Data Analysis

Partial Least Squares-Structural Equation Modeling was used for the analysis of results to analyse causal relationships between variables (Hair, Ringle and Sarstedt, 2011). This method was chosen due to its ability to handle complex models with multiple independent and dependent variables, even in small or medium-sized samples (Chin, 1998). Furthermore, unlike covariance-based SEM, PLS-SEM is more suitable when the primary goal is prediction and the exploration of emerging theoretical relationships (Sarstedt *et al.,* 2019). The partial least squares algorithm was used, employing an iterative approach to maximize the explained variance of the dependent variables. The model was run with 300 iterations and a 5% significance level for hypothesis testing, following the recommendations of Hair, Ringle and Sarstedt (2011). Considering the multidimensionality of social media marketing activity, it was measured as a second-order construct (Sarstedt *et al.*, 2019).

First-order measurement model

At the beginning of the process, EWOM 1 was eliminated due to its correlation with SMMA. Subsequently, the model was rerun to verify the reliability of each indicator. Its external loadings were evaluated considering the criterion that each must be greater than 0.7 (Hair, Ringle and Sarstedt, 2011). Cronbach's alpha and composite reliability scores were used to measure the internal consistency of the variables. The results surpassed the proposed threshold above 0.7 (Fornell and Larcker, 1981). In the same way, convergent validity was evaluated by the average variance extracted (AVE). A value greater than 0.5 indicates that at least 50% of the variance is explained by its indicators (Henseler, Ringle and Sarstedt, 2015). The results revealed values between 0.575 and 0.770, fulfilling this criterion. Table 2 presents the details of the reliability and convergent validity results of the first-order model.

Table 2: Reliability and convergent validity of the first-order model

Item	Loads	Variable	Cronbach's alpha	Composite reliability	AVE
ENT1	0,796	ENT	0,576	0,823	0,700
ENT2	0,875				
INT1	0,801	INT	0,536	0,811	0,682
INT2	0,850				
TRE1	0,809	TEN	0,583	0,826	0,704
TRE2	0,869				
PER1	0,842	PER	0,666	0,856	0,748
PER2	0,887				
ADV1	0,891	ADV	0,702	0,870	0,770
ADV2	0,864				
EWOM2	0,805	EWOM	0,702	0,834	0,626
EWOM3	0,803				
EWOM4	0,766				
IU1	0,757	IU	0,785	0,861	0,608
IU2	0,811				
IU3	0,801				
IU4	0,749				

The discriminant validity of the first-order model was admissible because the ratio between the Heterotrait–Monotrait (HTMT) correlations indicated in Table 3 was less than 0.9 (Henseler, Ringle and Sarstedt, 2015), confirming its consistency with the proposed parameters.

Table 3: HTMT

	IU	ENT	EWOM	INT	PER	ADV	TRE
IU							
ENT	0,798						
EWOM	0,814	0,716					
INT	0,747	0,693	0,866				
PER	0,762	0,830	0,811	0,852			
ADV	0,719	0,699	0,672	0,800	0,758		
TRE	0,829	0,879	0,870	0,829	0,862	0,799	

Second-order measurement model

The results of the evaluation of reliability, convergent and discriminant validity of the second-order mod was measured. In the case where only social media marketing, the results indicated that Cronbach's alpha (0.828) and composite reliability (0.879) demonstrated sufficient internal consistency. The HTMT criterion was used to evaluate the discriminant validity. As shown in table 4, the HTMT value of all variables was below 0.9, confirming that the discriminant validity of the model was admissible.

Table 4: Discriminant validity of the second-order model

	IU	EWOM	SMMA
IU			
EWOM	0,814		
SMMA	0,865	0,879	

Structural model analysis

The second stage of data analysis was evaluated through the structural model, with the objective of analysing the predictive capacity of the model and the relationships proposed in the study.

First, the collinearity of the exogenous constructs was excluded based on the values of the variance inflation factor, which were well below the threshold of 5. Therefore, no multicollinearity problem was observed. Regarding the analysis of the coefficients of determination, indicating the amount of variance explained by a construct (R^2), it can be observed that 52.2% of the variance of IU was explained by EWOM and SMMA. Furthermore, 45.4% of the variance of EWOM was explained by SMMA, with both values considered moderate (Chin, 1998).

The magnitude and statistical significance of the path coefficients, which specify the relationships of the structural model, were then assessed. To evaluate the significant effect, a test of bootstrapping of 5000 subsamples was performed with a p-value of 0.05 (Hair, Ringle and Sarstedt, 2011).

In relation to the first hypothesis, a positive and significant relationship was observed between SMMA and EWOM (β = 0.674; p-value = 0); thus, this hypothesis is statistically accepted. Regarding the second hypothesis, a positive and significant relationship was demonstrated between SMMA and IU (β = 0.530; p-value = 0), thereby accepting the hypothesis. In the fourth hypothesis, a positive relationship was observed between EWOM and UI; it was also significant (β = 0.250; p-value = 0.000). This confirmed the acceptance of this hypothesis. Finally, in relation to the third hypothesis, a positive relationship was observed in the mediating role of EWOM with SMMA and UI (β = 0.168; p-value = 0), confirming the approval of this hypothesis. Table 5 shows the hypotheses testing.

Table 5: Hypothesis test result

	Coefficient β	Statistics t	f^2	P-Values	Hypothesis
EWOM -> IU	0,250	3,958	0,071	0,000	Accepted
SMMA -> IU	0,530	9,705	0,321	0,000	Accepted
SMMA -> EWOM	0,674	17,569	0,833	0,000	Accepted
SMMA -> EWOM -> IU	0,168	3,830		0,000	Accepted

Multigroup analysis

Theoretically, multigroup analysis is the comparison of group-specific effects with the moderation of a variable (Baron and Kenny, 1986). Currently, this statistical technique is used to compare and contrast the relationships between variables from different groups and across different domains (Matthews, 2017). In addition, it allows researchers to test whether the relationships between variables are the same in different subgroups of a population (Yuan and Chan, 2016), affecting the direction and/or strength of the relationship between an independent or predictor variable. In line with this concept, group effects caused by the moderation of a variable express the degree of the group membership in each observation (Sarstedt, Henseler and Ringle, 2011). In this study, multigroup analysis was conducted using the permutation approach, which allows for the evaluation of significant differences between groups without requiring normal distributions (Chin and Dibbern, 2010). In the field of marketing, multigroup analysis is important for understanding and analysing the behaviour of various segments of one or more markets (Black and Babin, 2019).

Within the analysis, the comparison of relationships between the same variables based on educational level was evaluated. In the context of e-learning, the relationship between SMMA and EWOM considerably impacted users with a technical education level (β = 0.770), followed by university (β = 0.733) and postgraduate (β = 0.663) levels. Furthermore, the relationship between SMMA and UI was less significant for the postgraduate level of education (β = 0.460), being more significant for the university (β = 0.483) and technical (β = 0.483) educational levels. Similarly, the relationship between EWOM and UI, although to a lesser extent compared to the previous relationships, indicated its impact on users with university education (β = 0.292), followed by postgraduate (β = 0.308) and technical (β = 0.330) education levels. Table 6 shows the results of the multigroup analysis tests.

Table 6: Multigroup Analysis Result – Educational level

	β (Postgrad.)	Beta (Techn.)	Beta (University)	p-values (Postgrad.)	p-values (Techn.)	p-values (University)
SMMA→EWOM	0,663	0,770	0,733	0,000	0,000	0,000
SMMA → IU	0,460	0.483	0,483	0,000	0,000	0,000
EWOM → IU	0,308	0,330	0,292	0,007	0,002	0,010

5. Discussion of the Results

Within the context of e-learning, the study represents a contribution to the field of marketing. It evaluates the impact of social media marketing activities on electronic WOM generation and intention of use, offering a theoretical and practical approach that expands upon previous research on digital marketing and e-learning (Liaw, 2008; Ong and Lai, 2006; Roca, Chiu and Martínez, 2006; Tarhini, Hone and Liu, 2014). In view of the scarcity of studies in Latin America on e-learning, the present analysis is a contribution to the previous research and addresses a gap in the literature regarding the interactions between digital marketing and online education in this region (Azzari and Pelissari, 2020; Ramírez-Correa, Arenas-Gaitán and Rondán-Cataluña, 2015).

In this context, the study aims to enhance our understanding of the importance of the relationship between SMMAs in generating strong EWOM and its influence on students' UI. Differences were observed in the impact of specific digital marketing activities, with strategies that promote interactivity, personalization, and visually engaging content tending to generate a higher volume of positive EWOM. This supports and extends previous findings in digital marketing and consumer psychology (Aji, Nadhila and Sanny, 2020; Kim and Ko, 2010; Seo and Park, 2018; Sharma *et al.*, 2021) and aligns with findings from other authors in relation to the IU of users (Erkan and Evans, 2016; Goh, Heng and Lin, 2013; Knoll, 2016; Liaw, 2008). For instance, posts created by microlearning platforms on social media foster higher emotional engagement, a critical factor in the adoption of e-learning (Mujica, Villanueva and Lodeiros-Zubiria, 2021).

Establishing the relationship between SMMAs and EWOM enables companies to gain deeper insights into their audience and design more effective strategies for engaging users in educational contexts. This finding enhances the conceptual framework of digital marketing in e-learning, as users highly value interactive, accessible content that fosters an emotional connection with educational brands (Jalilvand and Samiei, 2012; Mujica, Villanueva and Lodeiros-Zubiria, 2021). Similarly, in the relationship between SMMAs and UI, students emphasized the importance of content generation in activities for creating an intention of use toward the brand (Knoll, 2016).

In addition, this research aims to fill a theoretical gap by demonstrating the mediation of EWOM between the SMMAs of e-learning platforms and the UI presented by students. The study concludes that EWOM serves as a pivotal mechanism for converting the positive perception of SMMAs into a stronger intention to use. This finding is consistent with existing literature that underscores the role of EWOM as a fundamental conduit between digital marketing strategies and consumer behavioural intentions (Lopes, 2011; Hasan, Al-Dmour and Al-Dmour, 2020; Jalilvand and Samiei, 2012), demonstrating the mediating role of EWOM is important for users to obtain genuine information from other users. This information allows users to recognize and confirm their expectations, in addition to resolving their doubts and forming their perceptions with less error rate in their intention to use (Cole, Shelley and Swartz, 2014; Hasan, Al-Dmour and Al-Dmour, 2020). This consistency aligns with previous research in the consumer marketplace (Jalilvand and Samiei, 2012) and consumer perceptions (Brunner, Stöcklin and Opwis, 2008).

Finally, with less significance in the results, it was possible to demonstrate the positive influence between EWOM and UI, as mentioned by several authors who indicated UI as one of the main consequences of generating EWOM (Jalilvand and Samiei, 2012; Pöyry *et al.*, 2012). This outcome allows students to reduce the risk when making a decision about their intention to use and allows companies to perform a predictive analysis of the intention to use of users, which is the basis for creating personalized marketing strategies for students (Wei *et al.*, 2021). Furthermore, this finding provides companies with a robust foundation for developing predictive models that incorporate the dynamics of EWOM, thereby enabling them to customize their digital marketing strategies to address the specific needs of students.

Multigroup analysis

In the multigroup analysis, the results have been classified according to the level of postgraduate, university, and technical education levels to measure their level of significance with the variables analysed in the study.

This indicates that, for postgraduate users, the relationships analysed between the variables were of lesser relevance. This is in comparison with technical users, who presented a higher level of significance among the relationships presented in the study. University users presented an intermediate level of significance in the relationship between the variables compared with technical and postgraduate users. The results and analysis of each relationship between the variables are described below:

The first relationship analysed between social media marketing activities and electronic WOM indicated greater significance among university students and professionals. This finding can be attributed to the professional and social development stage of these users, who are inclined to seek social validation and trusted sources of information when evaluating educational alternatives. This supports the idea that electronic WOM functions as a crucial mediator in educational decision-making, particularly in contexts where perceived risk is low due to minimal financial involvement. This suggesting that the information derived from electronic WOM, generated by the social media content of the institutions, was more relevant for them. It allowed them to evaluate course offerings and gain a genuine understanding of their target audience.

Similarly, for both educational levels, the relationship between social media marketing activities and the intention to use them was more important. This association may be attributed to the perceived accessibility, relevance, and personalization of the educational content delivered through digital platforms. Additionally, university and technical students tend to place greater importance on the institutions' ability to craft a coherent and compelling narrative that aligns with their professional aspirations. This suggests that the content generated by the institutions was more relevant. Moreover, they considered their communication in social media to be important, as it directly influenced the intention to use. This conclusion suggests that there was no fear of risk in the subsequent decision, likely attributed to the low monetary involvement associated with the educational service.

The level of technical education was more significant in the relationship between electronic WOM and intention to use, demonstrating the importance given to the information generated by other users. This can be explained by the fact that technical users, due to their limited prior experience with online education, are more dependent on electronic WOM to mitigate the uncertainty associated with the courses offered. This behaviour is consistent with technology adoption theories, such as the UTAUT2 model, which highlights the influence of social factors and performance expectations as critical determinants in the adoption decision (Tarhini *et al.,* 2017). This allowed them to obtain more honest and realistic information about the educational service and directly influenced their intention to use it in the future.

It also highlights a significant disparity in the relationship between variables for users with a postgraduate educational level, suggesting that the analysed e-learning method may not align with their expectations or their pursuit of online education. This inference may be attributed to their advanced level of education, as users with higher educational attainment are more likely to seek learning experiences that offer autonomy, competence, and a meaningful connection to their academic environment.. On the contrary, university students and technicians exhibited a higher affinity to study relations because they were in the process of continuing their professional growth and needed to strengthen their knowledge in a fast, adaptable, and inexpensive way. These features make the e-learning courses examined in this study attractive.

However, postgraduate users represented an intermediate level of significance between electronic WOM and the intention to use. This indicated the significance they attributed to the perception of other users in generating a subsequent intention to use. A pesar de su experiencia académica, este grupo demostró que confía en el WOM electrónico para validar las decisiones en entornos donde la calidad percibida de la educación en línea aún puede variar considerablemente. This highlights the necessity for institutions to adapt their marketing strategies to this demographic, focusing on attributes such as academic credibility, institutional reputation, and the perceived quality of content. This behaviour was driven by the authentic assessment users could obtain through their research of the platform in question.

6. Conclusion

The research provides us with an important perspective on the relationship between the variables analysed: social media marketing activities, electronic WOM, and intention to use, and the users of e-learning platforms. It also contributes to the analysis of the mediating role of electronic WOM in the context of e-learning platforms. Furthermore, this facilitates a more nuanced understanding of how these variables interact, offering valuable opportunities to optimize marketing strategies within the educational sector.

For the design of effective social media marketing campaigns, it is advisable to adopt strategies that prioritize the creation of engaging, relevant content that fosters user interaction. For instance, leveraging testimonials from successful students or positive experiences shared by current users can enhance the credibility of the content and facilitate positive eWOM. Moreover, implementing segmentation strategies based on users' educational level can improve the effectiveness of these campaigns, ensuring that the message is perceived as both relevant and personalized. Lastly, the development of free educational content series on platforms such as Instagram or TikTok can serve as a highly effective tactic for engaging younger audiences, who increasingly seek accessible and dynamic learning resources.In the multigroup analysis, the results have been classified based on the level of education, revealing greater importance among university students and technicians in the relationship between social media marketing activities and electronic WOM. This suggests that the information from electronic WOM generated by the social media content of the institutions was more relevant or them. It enabled them to evaluate course offerings and gain a genuine understanding of their target audience. Therefore, educational institutions can optimize their social media content by utilizing visual and narrative formats that highlight the practical and applied aspects of courses, which can be particularly appealing to technical and university students, as evidenced by the findings of this research.

In addition, the level of technical education significantly impacted the relationship between electronic WOM and the intention to use, reflecting the importance they attached to the information generated by other users. It allowed them to obtain more honest and realistic information about the educational service, directly influencing their intention to use it in the future. To maximize the impact of eWOM within this segment, it is recommended to create online communities where users can freely share their experiences. This will not only increase the volume of comments but also enhance trust in the platform.

It also highlights a significant disparity in the relationship between variables for users with a postgraduate educational level, suggesting that the analysed e-learning method may not align with their expectations or their pursuit of online education. This inference could be attributed to their already high level of education. On the contrary, university students and technicians exhibited a higher affinity to study relations because they were in the process of continuing their professional growth and needed to strengthen their knowledge in a fast, adaptable, and inexpensive way. These features make the e-learning courses examined in this study attractive. However, for postgraduate-level users, it would be prudent to explore marketing strategies that emphasize the depth and specialization of the courses, focusing on elements such as internationally recognized certification, opportunities for professional networking, and access to exclusive resources.

7. Limitations, and Suggestions for Future Research

The study has some limitations that should be considered for future research. First, the research focused on e-learning platforms with on-demand services. Accordingly, it is important to investigate other types of online educational services that allow enriching knowledge about the context to expand the information about the digital influence of these types of services.

Second, it would be beneficial to further investigate the dimensions of social media marketing activities and their relationship with each other in e-learning platforms to better understand how they may differentially influence the intention to use of users, including the mediating role of electronic WOM.

Third, the size of the sample analysed may be limited because it was reduced to residents of the metropolitan area of Lima. Therefore, future research could consider national and international participants, considering the cultural behaviour of each user.

Finally, the multigroup assessment was limited to the educational level of the sample. Accordingly, it would be interesting to examine other factors that represent a significant line of comparison, such as the influence of other more specific moderating variables, level of engagement, digital skills, and learning style through longitudinal studies to analyse how relationships evolve over time and whether differences exist between different life stages or educational contexts.

These recommendations can help to deepen the understanding of the effects of social media marketing in the context of e-learning and contribute to the development of more effective strategies in this area.

purposes. The study ensured strict adherence to ethical standards concerning human subjects, including the protection of participants' confidentiality and the anonymisation of all collected data.

AI Statement: In the preparation of this manuscript, the authors employed ChatGPT exclusively to enhance the linguistic quality and clarity of the text, as well as to assist in the translation process. The intellectual content, including the conceptual development, analysis, interpretation, and conclusions, was entirely conceived and authored by the researchers. No AI-generated content was utilised in the formulation or execution of the research itself.

Ethics Statement: Prior to their participation, all individuals were fully informed about the study's nature, objectives, and potential implications. Informed consent was obtained, with participants explicitly assured that their involvement was entirely voluntary and that their responses would be used solely for academic research

References

Aji, P. M., Nadhila, V. and Sanny, L. (2020) 'Effect of social media marketing on instagram towards purchase intention: Evidence from Indonesia's ready-to-drink tea industry', *International Journal of Data and Network Science*, 4(2), pp. 91–104. https://doi.org/10.5267/j.ijdns.2020.3.002.

Al-Shdayfat, N. M. (2018) 'Undergraduate student nurses' attitudes towards using social media websites: A study from Jordan', *Nurse Education Today*, 66, pp. 39–43. https://doi.org/10.1016/j.nedt.2018.03.017.

Alraimi, K. M., Zo, H. and Ciganek, A. P. (2015) 'Understanding the MOOCs continuance: The role of openness and reputation', *Computers and Education*, 80, pp. 28–38. https://doi.org/10.1016/j.compedu.2014.08.006.

Alrwashdeh, M., Emeagwali, O.L. and Aljuhmani, H.Y. (2019) 'The effect of electronic word of mouth communication on purchase intention and brand image: An applicant smartphone brands in North Cyprus', *Management Science Letters*, 9(4), pp. 505–518. https://doi.org/10.5267/j.msl.2019.1.011.

Anshari, M., Almunawar, M.N., Lim, S.A. and Al-Mudimigh, A. (2019) 'Customer relationship management and big data enabled: Personalization & customization of services', *Applied Computing and Informatics,* 15(2), pp. 94–101. https://doi.org/10.1016/j.aci.2018.05.004.

Azzari, V. and Pelissari, A. (2020) 'Does brand awareness influences purchase intention? The mediation role of brand equity dimensions', *Brazilian Business Review*, 17(6), pp. 669–685. https://doi.org/ 10.15728/BBR.2020.17.6.4.

Barbour, M.K., LaBonte, R., Kelly, K., Hodges, C., Moore, S., Lockee, B. et al. (2020) 'Understanding pandemic pedagogy: differences between emergency remote, remote, and online teaching', *State of the Nation: K-12 e-Learning in Canada*, (December), pp. 1–24. Available at: https://vtechworks.lib.vt.edu/handle/10919/101905

Baron, R. M. and Kenny, D. A. (1986) 'The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations.', *Journal of Personality and Social Psychology*, 51(6), pp. 1173–1182. https://doi.org/10.1037//0022-3514.51.6.1173.

Beaton, D.E., Bombardier, C., Guillemin, F. and Ferraz, M.B. (2000) 'Guidelines for the process of cross-cultural adaptation of self-report measures', *Spine*, 25(24), pp. 3186–3191. https://doi.org/10.1097/00007632-200012150-00014.

Bilgin, Y. (2018) 'The Effect of Social Media Marketing Activities on Brand Awareness, Brand Image and Brand Loyalty', *Business & Management Studies: An International Journal*, 6(1), pp. 128–148. https://doi.org/10.15295/bmij.v6i1.229.

Black, W. and Babin, B. J. (2019) 'Multivariate Data Analysis: Its Approach, Evolution, and Impact', *The Great Facilitator*, pp. 121–130. https://doi.org/10.1007/978-3-030-06031-2_16.

Breitsohl, J., Kunz, W. H. and Dowell, D. (2015) 'Does the host match the content? A taxonomical update on online consumption communities', *Journal of Marketing Management*, 31(9), pp. 1040–1064. https://doi.org/10.1080/0267257X.2015.1036102.

Brodie, R.J., Ilic, A., Juric, B. and Hollebeek, L. (2013) 'Consumer engagement in a virtual brand community: An exploratory analysis', *Journal of Business Research,* 66(1), pp. 105–114. https://doi.org/10.1016/j.jbusres.2011.07.029.

Bronnenberg, B. J., Kim, J. B. and Mela, C. F. (2016) 'Zooming in on choice: How do consumers search for cameras online?', *Marketing Science*, 35(5), pp. 693–712. https://doi.org/10.1287/mksc.2016.0977.

Brunner, T. A., Stöcklin, M. and Opwis, K. (2008) 'Satisfaction, image and loyalty: New versus experienced customers', *European Journal of Marketing*, 42(9–10), pp. 1095–1105. https://doi.org/10.1108/03090560810891163.

Carr, C. T. and Hayes, R. A. (2015) 'Social Media: Defining, Developing, and Divining', *Atlantic Journal of Communication*, 23(1), pp. 46–65. https://doi.org/10.1080/15456870.2015.972282.

Chan, T. K. H. *et al.* (2014) 'Antecedents and consequences of customer engagement in online brand communities', *Journal of Marketing Analytics*, 2(2), pp. 81–97. https://doi.org/10.1057/jma.2014.9.

Chatterjee, S. and Kumar Kar, A. (2020) 'Why do small and medium enterprises use social media marketing and what is the impact: Empirical insights from India', *International Journal of Information Management*, 53, p. 102103. https://doi.org/10.1016/j.ijinfomgt.2020.102103.

Chen, S. C. and Lin, C. P. (2019) 'Understanding the effect of social media marketing activities: The mediation of social identification, perceived value, and satisfaction', *Technological Forecasting and Social Change*, 140, pp. 22–32. https://doi.org/10.1016/j.techfore.2018.11.025.

Cheung, M. L., Pires, G. and Rosenberger, P. J. (2020) 'The influence of perceived social media marketing elements on consumer–brand engagement and brand knowledge', *Asia Pacific Journal of Marketing and Logistics*, 32(3), pp. 695–720. https://doi.org/10.1108/APJML-04-2019-0262.

Chin, W. W. (1998) 'Issues and opinion on structural equation modeling', *MIS Quarterly: Management Information Systems*, 22(1).

Chin, W.W. and Dibbern, J. (2010) "An introduction to a permutation based procedure for multi-group PLS analysis: Results of tests of differences on simulated data and a cross cultural analysis of the sourcing of information system services between Germany and the USA," in Handbook of Partial Least Squares. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 171–193.

Chu, S. C., Lien, C. H. and Cao, Y. (2019) 'Electronic word-of-mouth (eWOM) on WeChat: examining the influence of sense of belonging, need for self-enhancement, and consumer engagement on Chinese travellers' eWOM', *International Journal of Advertising*, 38(1), pp. 26–49. https://doi.org/10.1080/02650487.2018.1470917.

Cole, M. T., Shelley, D. J. and Swartz, L. B. (2014) 'Online instruction, e-learning, and student satisfaction: A three year study', *International Review of Research in Open and Distance Learning*, 15(6), pp. 111–131. https://doi.org/10.19173/irrodl.v15i6.1748.

Cook, D. A. and Triola, M. M. (2014) 'What is the role of e-learning? Looking past the hype', *Medical Education*, 48(9), pp. 930–937. https://doi.org/10.1111/medu.12484.

Düking, P., Holmberg, H. C. and Sperlich, B. (2018) 'The potential usefulness of virtual reality systems for athletes: A short SWOT analysis', *Frontiers in Physiology*, 9(MAR), pp. 1–4. https://doi.org/10.3389/fphys.2018.00128.

Eid, M. I. M. and Al-Jabri, I. M. (2016) 'Social networking, knowledge sharing, and student learning: The case of university students', *Computers and Education*. Elsevier Ltd, 99, pp. 14–27. https://doi.org/10.1016/j.compedu.2016.04.007.

Erkan, I. and Evans, C. (2016) 'The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption', *Computers in Human Behavior*, 61, pp. 47–55. https://doi.org/10.1016/j.chb.2016.03.003.

Etikan, I. (2016) "Comparison of convenience sampling and purposive sampling," *American journal of theoretical and applied statistics*, 5(1), p. 1. https://doi.org/10.11648/j.ajtas.20160501.11.

Evans, J. R. and Mathur, A. (2005) 'The value of online surveys', *Internet Research*, 15(2), pp. 195–219. https://doi.org/10.1108/10662240510590360.

Fornell, C. and Larcker, D. F. (1981) 'Evaluating Structural Equation Models with Unobservable Variables and Measurement Error', *Journal of Marketing Research*, 18(1), p. 39. https://doi.org/10.2307/3151312.

Ghewari, A. and Anute, N. (2021) 'Social Media Marketing Strategies Of E-Learning Organizations With Special Reference To Elementary Education', *Ilkogretim Online - Elementary Education Online*, 20(6), pp. 3593–3600. https://doi.org/10.17051/ilkonline.2021.06.339.

Godey, B., Manthiou, A., Pederzoli, D., Rokka, J., Aiello, G., Donvito, R. and Singh, R. (2016) 'Social media marketing efforts of luxury brands: Influence on brand equity and consumer behavior', *Journal of Business Research*, 69(12), pp. 5833–5841. https://doi.org/10.1016/j.jbusres.2016.04.181.

Goh, K. Y., Heng, C. S. and Lin, Z. (2013) 'Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content', *Information Systems Research*, 24(1), pp. 88–107. https://doi.org/10.1287/isre.1120.0469.

Goh, T.T., Yang, B., Dai, X. and Jin, D. (2017) 'A study of purchase influence and behavioral intention on the adoption of Electronic Word of Mouth (eWOM) systems', *Journal of Electronic Commerce in Organizations*, 15(3), pp. 14–32. https://doi.org/10.4018/JECO.2017070102.

Greenhow, C. and Lewin, C. (2016) 'Social media and education: reconceptualizing the boundaries of formal and informal learning', *Learning, Media and Technology*, 41(1), pp. 6–30. https://doi.org/10.1080/17439884.2015.1064954.

Hair, J. F., Ringle, C. M. and Sarstedt, M. (2011) 'PLS-SEM: Indeed a silver bullet', *Journal of Marketing Theory and Practice*, 19(2), pp. 139–152. https://doi.org/10.2753/MTP1069-6679190202.

Hasan, H., Al-Dmour, H. H. and Al-Dmour, R. H. (2020) 'The effect of ewom on e-loyalty: The mediating role of e-satisfaction', *International Journal of Customer Relationship Marketing and Management*, 11(3), pp. 82–100. https://doi.org/10.4018/IJCRMM.2020070105.

Hennig-Thurau, T., Gwinner, K.P., Walsh, G. and Gremler, D.D. (2004) 'Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet?', *Journal of Interactive Marketing*, 18(1), pp. 38–52. https://doi.org/10.1002/dir.10073.

Henseler, J., Ringle, C. M. and Sarstedt, M. (2015) 'A new criterion for assessing discriminant validity in variance-based structural equation modeling', *Journal of the Academy of Marketing Science*, 43(1), pp. 115–135. https://doi.org/10.1007/s11747-014-0403-8.

Hidayanto, A.N., Ovirza, M., Anggia, P., Budi, N.F.A. and Phusavat, K. (2017) 'The roles of electronic word of mouth and information searching in the promotion of a new E-commerce strategy: A case of online group buying in Indonesia', *Journal of Theoretical and Applied Electronic Commerce Research*, 12(3), pp. 69–85. https://doi.org/10.4067/S0718-18762017000300006.

Hoffman, D.L., Moreau, C.P., Stremersch, S. and Wedel, M. (2022) 'The Rise of New Technologies in Marketing: A Framework and Outlook', *Journal of Marketing*, 86(1), pp. 1–6. https://doi.org/10.1177/00222429211061636.

Ibrahim, B., Aljarah, A. and Ababneh, B. (2020) 'Do Social Media Marketing Activities Enhance Consumer Perception of Brands? A Meta-Analytic Examination', *Journal of Promotion Management*. Routledge, 26(4), pp. 544–568. https://doi.org/10.1080/10496491.2020.1719956.

Ibrahim, B., Aljarah, A. and Sawaftah, D. (2021) 'Linking social media marketing activities to revisit intention through brand trust and brand loyalty on the coffee shop Facebook pages: Exploring sequential mediation mechanism', *Sustainability*, 13(4), pp. 1–16. https://doi.org/10.3390/su13042277.

Jalilvand, M. R. and Samiei, N. (2012) 'The effect of electronic word of mouth on brand image and purchase intention: An empirical study in the automobile industry in Iran', *Marketing Intelligence and Planning*, 30(4), pp. 460–476. https://doi.org/10.1108/02634501211231946.

Kamboj, S., Sarmah, B., Gupta, S. and Dwivedi, Y. (2018) "Examining branding co-creation in brand communities on social media: Applying the paradigm of Stimulus-Organism-Response," *International Journal of Information Management*, 39, pp. 169–185. https://doi.org/10.1016/j.ijinfomgt.2017.12.001.

Kang, M.-J. (2024) 'A study on the impact of e-service quality in e-learning satisfaction', *Journal of Distribution Science*, 22(12), pp. 123–132. https://doi.org/10.15722/jds.22.12.202412.123.

Kaplan, A. and Haenlein, M. (2010) 'Users of the world, unite! The challenges and opportunities of Social Media', *Business Horizons*, 53(10), pp. 59–68. https://doi.org/10.1016/j.bushor.2009.09.003.

Kara, D. *et al.* (2018) 'The moderating effects of gender and income between leadership and quality of work life (QWL)', *International Journal of Contemporary Hospitality Management*, 30(3), pp. 1419–1435. https://doi.org/10.1108/IJCHM-09-2016-0514.

Katz, E., Blumler, J.G. and Gurevitch, M. (1973) "Uses and Gratifications Research," *Public Opinion Quarterly*, 37(4), p. 509. https://doi.org/10.1086/268109.

Khan, B. H. and Joshi, V. (2006) 'E-Learning Who, What and How?', *Journal of Creative Communications*, 1(1), pp. 61–74. https://doi.org/10.1177/097325860500100104.

Khan, M., Rubab, S., Awan, T. M., Khan, M., Malik, N., Daniyal, M., Ashraf, M. Z. and Kakar, A. S. (2022) 'The Relationship Between Social Media Marketing Activities and Brand Attachment: An Empirical Study from Pakistan', *The Journal of Asian Finance, Economics and Business*, 9(6), pp. 219–230. https://doi.org/10.13106/jafeb.2022.vol9.no6.0219.

Kim, H.L. and Hyun, S.S. (2019) "The relationships among perceived value, intention to use hashtags, eWOM, and brand loyalty of air travelers," *Sustainability*, 11(22), p. 6523. https://doi.org/10.3390/su11226523.

Kim, A. J. and Ko, E. (2010) 'Impacts of luxury fashion brand's social media marketing on customer relationship and purchase intention', *Journal of Global Fashion Marketing*, 1(3), pp. 164–171. https://doi.org/10.1080/20932685.2010.10593068.

Kim, A. J. and Ko, E. (2012) 'Do social media marketing activities enhance customer equity? An empirical study of luxury fashion brand', *Journal of Business Research*, 65(10), pp. 1480–1486. https://doi.org/10.1016/j.jbusres.2011.10.014.

Knoll, J. (2016) 'Advertising in social media: A review of empirical evidence', *International Journal of Advertising*, 35(2), pp. 266–300. https://doi.org/10.1080/02650487.2015.1021898.

Korhonen, V. (2024) *E-learning and digital education - Statistics & Facts*. Available at: https://www.statista.com/topics/3115/e-learning-and-digital-education/#topicOverview (Accessed: 30 May 2022).

Kramer, A. D. I., Guillory, J. E. and Hancock, J. T. (2014) 'Experimental evidence of massive-scale emotional contagion through social networks', *Proceedings of the National Academy of Sciences of the United States of America*, 111(24), pp. 8788–8790. https://doi.org/10.1073/pnas.1320040111.

Kumar, P. (2021) 'Enriching Blended Learning and Teaching Through the Use of Digital Marketing and Information Technology', *International Journal of Advanced Research in Education and Society*, 3(2), pp. 116–127. Available at: http://myjms.mohe.gov.my/index.php/ijares.

Lee, M. C. (2010) 'Explaining and predicting users' continuance intention toward e-learning: An extension of the expectation-confirmation model', *Computers and Education*, 54(2), pp. 506–516. https://doi.org/10.1016/j.compedu.2009.09.002.

Lee, S.-J. and Shin, S.-Y. (2014) 'The Effect of Purchase Reviews of Internet Shopping mall on Benefits Sought of Sales Promotion, Fashion Customer's Purchase Satisfaction, Repurchase Intention, and Word-of-Mouth Intention', *Fashion & Textile Research Journal*, 16(1), pp. 79–90. https://doi.org/10.5805/sfti.2014.16.1.79.

Li, F., Larimo, J. and Leonidou, L. C. (2021) 'Social media marketing strategy: definition, conceptualization, taxonomy, validation, and future agenda', *Journal of the Academy of Marketing Science*, 49(1), pp. 51–70. https://doi.org/10.1007/s11747-020-00733-3.

Liaw, S. S. (2008) 'Investigating students' perceived satisfaction, behavioral intention, and effectiveness of e-learning: A case study of the Blackboard system', *Computers and Education*, 51(2), pp. 864–873. https://doi.org/10.1016/j.compedu.2007.09.005.

Lin, H.-C., Han, X., Lyu, T., Ho, W.-H., Xu, Y., Hsieh, T.-C., Zhu, L. and Zhang, L. (2020) 'Task-technology fit analysis of social media use for marketing in the tourism and hospitality industry: a systematic literature review', *International Journal of Contemporary Hospitality Management*, 32(8), pp. 2677–2715. https://doi.org/10.1108/IJCHM-12-2019-1031.

Liu, S. H., Liao, H. L. and Pratt, J. A. (2009) 'Impact of media richness and flow on e-learning technology acceptance', *Computers and Education*, 52(3), pp. 599–607. https://doi.org/10.1016/j.compedu.2008.11.002.

Liu, X., Shin, H. and Burns, A. C. (2021) 'Examining the impact of luxury brand's social media marketing on customer engagement: Using big data analytics and natural language processing', *Journal of Business Research*, 125, pp. 815–826. https://doi.org/10.1016/j.jbusres.2019.04.042.

Lopes, S. D. F. (2011) 'Destination image: Origins, Developments and Implications', *PASOS. Revista de Turismo y Patrimonio Cultural*, 9(2), pp. 305–315. https://doi.org/10.25145/j.pasos.2011.09.027.

de Matos, C. A. and Rossi, C. A. V. (2008) 'Word-of-mouth communications in marketing: A meta-analytic review of the antecedents and moderators', *Journal of the Academy of Marketing Science*, 36(4), pp. 578–596. https://doi.org/10.1007/s11747-008-0121-1.

Matthews, L. (2017) 'Applying multigroup analysis in PLS-SEM: A step-by-step process', *Partial Least Squares Path Modeling: Basic Concepts, Methodological Issues and Applications*. https://doi.org/10.1007/978-3-319-64069-3_10.

Merrilees, B. (2016) 'Interactive brand experience pathways to customer-brand engagement and value co-creation', *Journal of Product and Brand Management*, 25(5), pp. 402–408. https://doi.org/10.1108/JPBM-04-2016-1151.

Mujica, A., Villanueva, E. and Lodeiros-Zubiria, M.L. (2021) "Micro-learning platforms Brand Awareness using socialmedia marketing and Customer Brand Engagement," *International Journal of Emerging Technologies in Learning*, 16(17), p. 19. https://doi.org/10.3991/ijet.v16i17.23339.

Mukhtar, K. *et al.* (2020) 'Advantages, limitations and recommendations for online learning during covid-19 pandemic era', *Pakistan Journal of Medical Sciences*, 36(COVID19-S4), pp. S27–S31. https://doi.org/10.12669/pjms.36.COVID19-S4.2785.

Naaman, M., Becker, H. and Gravano, L. (2011) 'Hip and trendy: Characterizing emerging trends on Twitter', *Journal of the American Society for Information Science and Technology*, 62(5), pp. 902–918. https://doi.org/10.1002/ASI.21489.

Noh, J. S., Jang, H. Y. and Jeon, S. S. (2021) 'The influence of e-WOM information characteristics on learning trust and e-WOM intention among online learning', *Journal of Logistics, Informatics and Service Science*, 8(2), pp. 134–150. https://doi.org/10.33168/LISS.2021.0208.

Ong, C. S. and Lai, J. Y. (2006) 'Gender differences in perceptions and relationships among dominants of e-learning acceptance', *Computers in Human Behavior*, 22(5), pp. 816–829. https://doi.org/10.1016/j.chb.2004.03.006.

Panigyrakis, G., Panopoulos, A. and Koronaki, E. (2020) 'All we have is words: applying rhetoric to examine how social media marketing activities strengthen the connection between the brand and the self', *International Journal of Advertising*, 39(5), pp. 699–718. https://doi.org/10.1080/02650487.2019.1663029.

Pöyry, E., Parvinen, P., Salo, J. y Blakaj, H. (2012) 'Hedonic and utilitarian search for electronic word-of-mouth', *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 1797–1806. https://doi.org/10.1109/HICSS.2012.300.

Ramírez-Correa, P. E., Arenas-Gaitán, J. and Rondán-Cataluña, F. J. (2015) 'Gender and acceptance of e-learning: A multi-group analysis based on a structural equation model among college students in Chile and Spain', *PLoS ONE*, 10(10). https://doi.org/10.1371/journal.pone.0140460.

Roca, J. C., Chiu, C. M. and Martínez, F. J. (2006) 'Understanding e-learning continuance intention: An extension of the Technology Acceptance Model', *International Journal of Human Computer Studies*, 64(8), pp. 683–696. https://doi.org/10.1016/j.ijhcs.2006.01.003.

Rusfian, E. Z. and Alessandro, J. D. (2021) 'Influence of Social Media'S Marketing Activity on Local Brand Equity and Consumer Response: Using Mix Method Approach', *Linguistics and Culture Review*, 5(S1), pp. 767–780. https://doi.org/10.21744/lingcure.v5ns1.1462.

Salloum, S. A. and Shaalan, K. (2019) 'Factors Affecting Students' Acceptance of E-Learning System in Higher Education Using UTAUT and Structural Equation Modeling Approaches', *Advances in Intelligent Systems and Computing*. Springer International Publishing. https://doi.org/10.1007/978-3-319-99010-1_43.

Sarstedt, M., Hair Jr, J. F., Cheah, J.-H., Becker, J.-M. y Ringle, C. M. (2019) 'How to specify, estimate, and validate higher-order constructs in PLS-SEM', *Australasian Marketing Journal*, 27(3), pp. 197–211. https://doi.org/10.1016/j.ausmj.2019.05.003.

Sarstedt, M., Henseler, J. and Ringle, C. M. (2011) 'Multigroup analysis in partial least squares (PLS) path modeling: Alternative methods and empirical results', *Advances in International Marketing*, 22(2011), pp. 195–218. https://doi.org/10.1108/S1474-7979(2011)0000022012.

Seo, E. J. and Park, J. W. (2018) 'A study on the effects of social media marketing activities on brand equity and customer response in the airline industry', *Journal of Air Transport Management*, 66, pp. 36–41. https://doi.org/10.1016/J.JAIRTRAMAN.2017.09.014.

Shams, M. S., Niazi, M. M., Gul, H., Mei, T. S. y Khan, K. U. (2022) 'E-Learning Adoption in Higher Education Institutions During the COVID-19 Pandemic: A Multigroup Analysis', *Frontiers in Education*, 6. https://doi.org/10.3389/feduc.2021.783087.

Shanahan, T., Tran, T. P. and Taylor, E. C. (2019) 'Getting to know you: Social media personalization as a means of enhancing brand loyalty and perceived quality', *Journal of Retailing and Consumer Services*, 47, pp. 57–65. https://doi.org/10.1016/j.jretconser.2018.10.007.

Sharma, S., Singh, S., Kujur, F. y Das, G. (2021) 'Social media activities and its influence on customer-brand relationship: An empirical study of apparel retailers' activity in india', *Journal of Theoretical and Applied Electronic Commerce Research*, 16(4), pp. 602–617. https://doi.org/10.3390/jtaer16040036.

Shehzadi, S., Nisar, Q. A., Hussain, M. S., Basheer, M. F., Hameed, W. U. y Chaudhry, N. I. (2021) 'The role of digital learning toward students' satisfaction and university brand image at educational institutes of Pakistan: a post-effect of COVID-19', *Asian Education and Development Studies*, 10(2), pp. 276–294. https://doi.org/10.1108/AEDS-04-2020-0063.

Stafford, M. R., Himelboim, I., Walter, D. y Ophir, Y. (2022) "The evolution of advertising research through four decades: a computational analysis of themes, topics and methods," *International Journal of Advertising*, 42(1), pp. 1–24. https://doi.org/10.1080/02650487.2022.2128005.

Stanley, D. and Montero Fortunato, Y. R. (2022) 'The efficacy of online higher education in Latin America: A systematic literature review', *Revista Iberoamericana de Tecnologías del Aprendizaje: IEEE-RITA*, 17(3), pp. 262–269.

Stephen, A. T. and Galak, J. (2012) 'The effects of traditional and social earned media on sales: A study of a microlending marketplace', *Journal of Marketing Research*, 49(5), pp. 624–639. https://doi.org/10.1509/jmr.09.0401.

Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y. y Yeh, D. (2008) "What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction," *Computers & Education*, 50(4), pp. 1183–1202. https://doi.org/10.1016/j.compedu.2006.11.007.

Talebian, S., Mohammadi, H. M. and Rezvanfar, A. (2014) 'Information and Communication Technology (ICT) in Higher Education: Advantages, Disadvantages, Conveniences and Limitations of Applying E-learning to Agricultural Students in Iran', *Procedia - Social and Behavioral Sciences*, 152, pp. 300–305. https://doi.org/10.1016/j.sbspro.2014.09.199.

Tan, W. *et al.* (2014) 'A Trust Evaluation Model for E-Learning Systems', *Systems Research and Behavioral Science*, 31(3), pp. 353–365. https://doi.org/10.1002/sres.2283.

Tarhini, A., Hone, K. and Liu, X. (2014) 'The effects of individual differences on e-learning users' behaviour in developing countries: A structural equation model', *Computers in Human Behavior*, 41, pp. 153–163. https://doi.org/10.1016/j.chb.2014.09.020.

Tran, T. P., Muldrow, A. and Ho, K. N. B. (2021) 'Understanding drivers of brand love - the role of personalized ads on social media', *Journal of Consumer Marketing*, 38(1), pp. 1–14. https://doi.org/10.1108/JCM-07-2019-3304.

Troussas, C., Krouska, A. and Sgouropoulou, C. (2021) 'Impact of social networking for advancing learners' knowledge in E-learning environments', *Education and Information Technologies*, 26(4), pp. 4285–4305. https://doi.org/10.1007/s10639-021-10483-6.

Vate-U-Lan, P. and Masouras, P. (2018) 'Thriving social network for communication on elearning: Exploring gender differences in attitudes', *ACM International Conference Proceeding Series*, pp. 1–6. https://doi.org/10.1145/3279996.3280010.

Wang, W., Liu, H., Wu, Y. J. y Goh, M. (2023) "Disconfirmation effect on online reviews and learner satisfaction determinants in MOOCs," *Education and Information Technologies*, 28(12), pp. 15497–15521. https://doi.org/10.1007/s10639-023-11824-3.

Wei, G., Wang, L., Lin, W., Yanxiong, W. y Musse, S. Y. (2021) 'The relationship of sustainability communication on social media with banking consumers' loyalty through e-WOM', *Sustainability*, 13(7), pp. 1–13. https://doi.org/10.3390/su13073832.

Wibowo, A., Chen, S.-C., Wiangin, U., Ma, Y. y Ruangkanjanases, A. (2021) 'Customer behavior as an outcome of social media marketing: The role of social media marketing activity and customer experience', *Sustainability*, 13(1), pp. 1–18. https://doi.org/10.3390/su13010189.

Xu, H., Lovett, J.C. and Law, R. (2022) "Understanding destination value co-creation on social media: An application of travel blog analysis," *Tourism and Hospitality*, 3(3), pp. 573–588.https://doi.org/10.3390/tourhosp3030035.

Yadav, M. and Rahman, Z. (2017) 'Measuring consumer perception of social media marketing activities in e-commerce industry: Scale development & validation', *Telematics and Informatics*, 34(7), pp. 1294–1307. https://doi.org/10.1016/J.TELE.2017.06.001.

Yadav, M. and Rahman, Z. (2018) 'The influence of social media marketing activities on customer loyalty: A study of e-commerce industry', *Benchmarking*, 25(9), pp. 3882–3905. https://doi.org/10.1108/BIJ-05-2017-0092.

Yuan, K. H. and Chan, W. (2016) 'Measurement invariance via multigroup SEM: Issues and solutions with chi-square-difference tests', *Psychological Methods*, 21(3), pp. 405–426. https://doi.org/10.1037/met0000080.

Zandi, G., Lahrash, H. A. E. and Shakhim, F. R. (2022) 'Factors affecting the adoption of e-learning: An empirical study of Libyan universities', *Journal of Information Technology Management*, 14(4), pp. 95–117. https://doi.org/10.22059/jitm.2022.88292.

Zhang, J., She, L., Wang, D. y Shafiq, A. (2022) "Chinese consumers' E-learning satisfaction and continuance purchase intention on paid online Python course," *Frontiers in Psychology*, 13, p. 849627. https://doi.org/10.3389/fpsyg.2022.849627.

Zhao, Y., Liu, Y., Lai, I. K. W., Zhang, H. y Zhang, Y. (2016) 'The impacts of attitudes and engagement on electronicword of mouth (eWOM) of mobile sensor computing applications', Sensors, 16(3). https://doi.org/10.3390/s16030391.

Appendix 1: Questionnaire

Social Media Marketing Activities (SMMA) (Aji et al., 2020)

Dimensions: Entertainment (ENT), Interaction (INT), Trend (TEN), Personalization (PER), and Advertising (PUB).

Entertainment (ENT)

ENT1: Interacting with the social media platforms of the e-learning platform is fun.

ENT2: The content of the e-learning platform on its social media is interesting to me.

Interaction (INT)

INT1: The e-learning platform's social media allows me to share information with others.

INT2: It is easy for me to express my opinion through social media about the e-learning platform.

Trend (TEN)

TRE1: The content shared by the e-learning platform on its social media appears to be in line with the current trend.

TRE2: Interacting with the social media platforms of the e-learning platform is what is commonly done nowadays.

Personalization (PER)

PER1: The e-learning platform's social media provides me with the information I need.

PER2: The e-learning platform's social media provides me with the information I need.

Advertising (ADV)

ADV1: I like the advertising on social media that the e-learning platform has posted; it catches my attention.

ADV2: I like the advertising on social media that the e-learning platform has posted; it catches my attention.

Electronic Word of Mouth (eWOM) (Hidayanto, et al., 2017)

eWOM1: Online course recommendations are useful to me.

eWOM2: Recommendations regarding online courses influence my choice of e-learning platform.

eWOM3: Recommendations regarding online courses would increase my interest in obtaining more information.

eWOM4: I can make the decision to choose a specific e-learning platform based on a recommendation I received.

Intention to Use (IU) (Alraimi, et al., 2015)

IU1: I intend to continue using the e-learning platform in the future.

IU2: I will continue using the e-learning platform in the future.

IU3: I will fully recommend the e-learning platform for others to use.

IU4: I will continue using the e-learning platform as regularly as I do now.

The Video Lecture as an Instructional Method of MOOCS: Impact on the Students' Re-Watching Behaviors

Muhammed Shittu[1], Anna Gorbunova[2], Christopher Lange[3] and Jamie Costley[1]

[1]College of Education, United Arab Emirates University, Al Ain, United Arab Emirates

[2]Institute of Education, HSE University, Moscow, Russia

[3]Dankook University, Department of British and American Humanities, Gyeonggi-do, South Korea

202170076@uaeu.ac.ae
agorbunova@hse.ru
christopherhlange@dankook.ac.kr
jcostley@uaeu.ac.ae (corresponding author)

https://doi.org/10.34190/ejel.23.2.3587

Abstract: This study examines the relationship between instructional delivery that affects students' engagement with video lectures and the need to re-watch video lectures in the context of Massive Open Online Courses (MOOCs), where video lectures are the predominant mode of instruction. Multiple linear regression was used to investigate the relationships between some key characteristics of video lectures— the arrangement of text and graphs that creates eye scanning behavior, audiovisual support, text-image comprehensibility, professor's speech clarity—with the need to re-watch video lectures. The findings from the study reveal that the arrangement of text and graphs, the professor's speech clarity, and text-image comprehensibility emerge as major factors influencing the need to re-watch video lectures in MOOCs. However, audiovisual support does not significantly influence the need to re-watch video lectures. Consequently, the study sheds light on the positive relationship between the arrangement of text and graphs and the need to re-watch video lectures, highlighting the potential cognitive overload experienced by students when oscillating between textual and graphical information. Similarly, the study also shows how the comprehensibility of text and images impacts students' video lecture rewatching behavior, emphasizing the necessity for clear and comprehensible visuals in instructional materials. Additionally, our findings indicate the significance of speech clarity in video lectures, with clear speech patterns contributing to a reduced need for re-watching video lectures. However, the study notes the limited impact of audiovisual support on students' re-watching behavior, suggesting the need for instructors and content creators to minimize audiovisual distractions in order to create sustained engagement. While previous studies have explored various aspects of student engagement in MOOCs, this research fills a critical gap by specifically investigating instructional delivery affecting engagement behaviors of students and the propensity for such instructional delivery to cause students to re-watch video lectures in Moocs. Thus, these findings highlight key areas for optimizing video lecture design in MOOCs, offering actionable insights for educators and content creators to enhance instructional effectiveness and student engagement in fully online learning environments.

Keywords: Student engagement, Video lectures, Online learning, MOOCs

1. Introduction

Online learning is becoming more prevalent in education, and it comes in a variety of forms, including virtual learning environments (VLEs), Massive Open Online Courses (MOOCs), and learning management systems (LMS) (Bonafini, 2017). It can be referred to as learning experiences that occur in asynchronous or synchronous environments using a variety of devices that can be connected to the internet, enabling students to learn from anywhere, anytime (Cojocariu, et al., 2014; Dhawan, 2020). One way to think of online learning is as an instrument that can help with teaching and learning processes, making learning more adaptable, dynamic, and centered on the needs of the student. The benefits of online learning extend to a broader base of students with similar goals through the use of MOOCs. According to Zhang, Gao and Zhang (2021), MOOCs are communication and collaboration platforms that allow users to share and strengthen their knowledge. MOOCs have gained traction in higher education and have become significant components of the education process (Jin and Shang, 2022). Numerous stakeholders, including students, academics, and corporate professionals have reportedly benefited from MOOCs (Egloffstein and Ifenthaler, 2017). Furthermore, more educational institutions have started embracing MOOCs, which have revolutionized the higher education landscape by providing unprecedented access to a wide range of learning resources through a variety of models (Burd, Smith and Reisman 2015).

Video lectures play a key role in the online learning process, particularly in MOOCs where they are the main form of instruction and student engagement (Atapattu and Falkner, 2018). Overall, video lectures have been

Cite this article: Shittu, M. et al. 2025. “The Video Lecture as an Instructional Method of MOOCS: Impact on the Students' Re-Watching Behaviors”, *Electronic Journal of e-Learning*, 23(2), pp 64-79, https://doi.org/10.34190/ejel.23.2.3587

reported to be helpful in the preparation and facilitation of learning by students (Roehling, 2018). They are considered capable of engaging learners in ways that allow them to take control of their learning processes, giving them access to a number of learner-control options in which the learner can choose how to engage with the video, including the ability to segment, pause, re-watch, and control the speed of the video based on individual learning needs (Hwang, Lai and Wang, 2015). In terms of extant research regarding those learner-control options, conflicting results have been found with rewatching videos in particular. Recently, Palmer, Chu and Persky (2019) and Patel, et al. (2019) demonstrated that students who rewatched video lectures had better performance compared with students who did not rewatch video lectures. In contrast, Martin, Mills, D'Mello, and Risko (2018) revealed that there are no benefits of rewatching video lectures on the retention of lecture content. Another study also noted that the ability to rewatch video lectures can lead to a less engaged learning experience (Guo, Kim and Rubin, 2014). Research has shown that students who frequently rewatch or skip around in video lectures are more likely to experience disengagement and reduced learning outcomes (Kim, et al., 2014).

The way in which instruction is delivered through video lectures may ultimately play a role in how learners engage with the videos, particularly in their decision whether to re-watch videos. Research has reported that one of the main reasons learners feel the need to re-watch video lectures is due to their lack of comprehension during the initial viewing of a video lecture (Le, et al., 2010). Aside from the natural complexity of the content, there are number of purely instructional factors (replaying brief segment, repeating non visual explanation, returning to missed content) that contribute to a lack of student comprehension in regard to the content delivered via video lectures (Kim, et al., 2014), including the inefficient arrangement of textual and graphical information, distracting audiovisuals, ineffective use of text and images to support instruction, and the lack of instructor's speech clarity within the video lecture. However, there is a lack of research investigating the direct relationship between those instructional factors and the need to re-watch video lectures. This study seeks to address this gap by delving into the intricate dynamics of student engagement with video lecture content in MOOCs by using multiple linear regression to examine the relationships between specific behaviors and re-watching videos. These behaviors include scanning between texts and graphs, processing of distracting audio-visual information, utilizing text and images for more efficient processing, and listening to instruction directly delivered from the instructor, which may ultimately influence their decision to re-watch video lectures. .

2. Literature Review

While the video lectures are a prevalent means of instruction in online learning, its optimal instructional design features are often an open question, leading to ongoing debate and uncertainty within the field. This literature review gives an overview of the current body of research investigating the influence of diverse instructional design elements, including content delivery, audiovisual support, instructor clarity, and text-image comprehensibility, on student engagement, cognitive load and the frequency of video lecture re-watching in MOOCs.

2.1 Importance of Video Lectures in MOOCs

MOOCs have been expanding rapidly, and this initiative has received support from numerous prestigious universities and organizations (Albelbisi, Al-Adwan and Habibi, 2023). More than 23 million learners enrolled in at least one MOOC in 2016, bringing the total number of learners to 58 million. There are 6,850 MOOC courses offered by more than 700 universities (Shah, 2018). Different studies have shown that MOOCs have become more and more important in learning and have become increasingly popular in higher education (Aparicio, et al., 2019; Pozón-López, et al., 2020; Tzeng, et al., 2022).

Video-based learning materials are increasingly used in online learning environments around the world (Akcapinar and Bayazit, 2018). The prevalence of video lectures in contemporary education, particularly in hybrid or fully online classes, has surged, marking a significant shift in teaching (Scagnoli, McKinney and Moore-Reynen, 2015). Several research studies have demonstrated that the use of video lectures can be a highly effective instrument for instruction (Rackaway, 2012; Hsin and Cigas, 2013; Stockwell, et al., 2015). Notably, the utilization of video lectures has demonstrated a considerable improvement in student engagement compared to traditional text-based course materials (June, Yaacob and Kheng, 2014), leading to enhanced retention and reduced teacher intervention (Hsin and Cigas, 2013).

With the prevalence of computers and mobile devices and the increase of MOOCs, more learners are favoring video-based lectures over traditional lectures (Banoor, Issack and Frank, 2019). The video lectures are a primary teaching method within MOOCs (Atapattu and Falkner, 2018), therefore understanding its importance is crucial for unravelling the dynamics between instructional factors affecting student engagement and the subsequent need to re-watch instructional video lectures in an online learning environment. Furthermore, studies revealed that students spend more time watching videos than reading texts (Seaton, et al., 2014). A video lecture within the context of a MOOC constitutes a composite integration of various elements, encompassing instructor-led lectures, quizzes, and presentation slides (Sinha, Jermann and Dillenbourg,., 2014).

One of the main advantages of video lectures is the visual and auditory nature of the content, which enhances the accessibility of complex topics, making education available to individuals who might face barriers in traditional learning environments (Akcapinar and Bayazit, 2018). Besides, it provides students the opportunity to advance in their learning autonomously, enabling them to revisit specific sections at their discretion—a feature that can be delineated as among other noteworthy advantages (Kim, et al., 2014). Furthermore, MOOCs offer students flexibility in the learning process by enabling them to pause, repeat, or skip video lectures (Triay, et al., 2016).

The role of video lectures has been well explored, for example Lange and Costley (2020) studied the challenges and issues arising from media use in online video lectures. The research identifies five primary categories of media delivery problems: pace, intelligibility, quality, media diversity, and congruence. The study examines a random sample of OCU video lectures across various disciplines, revealing how these media-related issues can hinder the learning process. The results emphasize the need for instructors to carefully design and deliver media in online lectures to enhance the learning experience. Instructors can do this by giving autonomy to learners through learner control while at the same time providing access to instructor-learner communication where feedback can efficiently be given outside of the video lectures. For example, results of one another study indicate that perceived active control, perceived synchronicity, and perceived two-way communication significantly influence both the participants' intention to complete the MOOC and engagement on the platform (Shao & Chen, 2021).

Another study of Brame (2016) also explores the effective use of educational videos in higher education, emphasizing their role in flipped, blended, and online classes. The study found that three key principles, cognitive load management, student engagement maximization, and active learning promotion, form the foundation for effective video use. Koh and Ahn (2023) examine the role of video lecture types in motivating students within sustainable flipped learning environments. Using a mixed-methods design, it compared traditional video lectures (TVL) created by instructors with student-engaged video lectures (SEVL), co-created by instructors and students. Findings revealed no significant overall difference in motivation between TVL and SEVL, though SEVL slightly improved attention regardless of student grades.

The study of Kuznekoff (2020) found that while viewing more content positively influenced learning, the effect plateaued with excessive viewing. Only 34% of students completed full lectures, with the majority 60%, either partially watched or disengaged entirely. These findings highlight the need to optimize video length and format to enhance engagement and learning outcomes in online education.

Abakumov, et al. (2018) studied the effect of rewatching video lectures in MOOCs; they use a cross-classification multilevel logistic approach to analyze data from four Coursera courses in social sciences and math. The results revealed that the overall effect of rewatching video lectures on assessment performance varies significantly—sometimes improving, having no impact, or even negatively affecting performance, depending on the course and specific assessment items. This nuanced finding suggests that the effectiveness of rewatching video lectures is not uniform across all students or assessment types, challenging the general recommendation to rewatch videos as a preparation strategy. The study underscores the need for further research to understand the factors contributing to these varying effects, with implications for tailoring instructional strategies to individual learner needs.

2.2 The Impact of Content Delivery on Student Engagement with Video Lectures

The concept of student engagement in online learning environments is multifaceted, encompassing various dimensions that contribute to the overall learning experience. Student engagement is widely recognized as crucial to learning and satisfaction in online courses (Maloney, et al., 2023). Student engagement refers to the level of involvement and active participation that students demonstrate in their online courses (Shao and Chen, 2021). Student engagement can also be said to be associated with positive educational outcomes, such as enhanced academic achievement, improved retention and persistence rates, and increased motivation and satisfaction within the classroom setting (Maloney, et al., 2023). In the context of massive open online courses, student engagement with video lectures is a critical component of the learning experience. While video lectures are a common feature of MOOCs, research has shown that students may disengage from these materials for various reasons, such as fatigue, loss of motivation, or challenges in effectively processing the content (Li, et al., 2015). Understanding the factors that contribute to student engagement with video lectures in MOOCs can provide valuable insights into the design and delivery of online learning experiences that better meet the needs of diverse learners. For example, Guo, Kim and Rubin (2014) found that shorter videos (less than 6 minutes) were found to be significantly more engaging. Videos that combine an instructor's "talking head" with slides surpass the engagement levels of slides alone.

One critical aspect of students' engagement in MOOCs is the level of interaction with video lectures, which serve as a primary mode of content delivery (Guo, Kim and Rubin, 2014). Although behavioral engagement usually implies watching the video itself, it goes beyond mere passive viewing; it involves active participation, thoughtful reflection, and the application of acquired knowledge (Bonafini, 2017). The impact of student engagement, a critical aspect of effective learning, is intricately connected to various factors related to the incorporation of video lectures. The types of videos used, the organizational integration of videos in teaching methodologies, and the adaptive strategies employed by instructors to accommodate videos all influence student engagement significantly (Guo, Kim and Rubin, 2014). Studies also suggest that supplementary content interaction associated with video lectures can lead to further engagement, ultimately correlating with completion rates. For example, active participation in discussion forums emerges as a strong predictor of MOOC completion, aligning with prior literature on the significance of forum engagement (Bonafini, 2017). Results of another study show a strong correlation between such learning activities as forum engagement and content interaction and student engagement levels (Hussain, et al., 2018). A significant number of studies have explored engagement as it relates to video lectures, for example Costley, Hughes and Lange (2017) investigated the relationship between instructional design and student engagement with video lectures at a cyber-university in South Korea. The study sought to identify how various aspects of instructional design influence the likelihood of students watching and completing video lectures. The study found significantly higher rates of students both watching and completing their video lectures. The findings revealed positive correlations between student engagement and five key instructional design elements: designing methods, setting the curriculum, establishing time parameters, enforcing netiquette, and effectively utilizing the medium. This aids learners and prevents lack of comprehension often associated with video lectures (Le, et al., 2010).

Lackmann, et al. (2021) studied the influence of video format on engagement and performance in online learning. The study compares infographic videos (animated graphics and text) to lecture capture (professor delivering a classroom lecture) and found distinct impacts on engagement and learning outcomes. Lecture capture elicited higher emotional engagement in the short term, while infographic videos sustained cognitive and emotional engagement over longer periods. The study highlights that optimal video formats depend on the desired balance between engagement and learning outcomes.

Cognitive Load Theory and its Effects on Video Lecture

During the design and production process of multimedia instructional materials, including video lectures, cognitive load is considered one of the most important considerations (Afify, 2020). The cognitive load theory proposes that the human brain can process only a finite amount of information and that effective learning occurs when the cognitive load reflects a person's cognitive capabilities (Sweller, 2022). In other words, Sweller (2024) defines cognitive load as the amount of information stored and processed in working memory. This theory is significant in the context of video lectures, where students may be exposed to a large amount of visual and auditory information, thus overloading their cognitive capacities. Research has shown that cognitive load can have an enormous effect on the effectiveness of video lectures in Massive Open Online Courses (Brame, 2016; Costley, et al., 2021; Afify 2020). Specifically, the way in which information is presented in video lectures can impact the cognitive load experienced by students, which in turn may influence their engagement with video lectures, Chen and Wu (2015) found that students who were presented with a voice-over style of video lecture experienced significantly lower levels of cognitive load than those who experienced a visualizer style. Another study found that presenting content in video lectures that prevents learners from processing information unrelated to learning can lead to higher levels of engagement (Lange & Costley, 2020). This suggests that the design of video lectures should prioritize minimizing cognitive load in order to enhance student engagement with video lectures which in turn may influence the need to re-watch the lectures.

2.3 Arrangement of Text and Graphs

By using various formats to deliver information, the application of visual media has been noted to assist students in improving learning outcomes in online learning platforms (Lange and Costley, 2020). The arrangement of text and graphs in an online learning platform is a key factor in maximizing the comprehension and learning outcomes of the learners. In addition, the layout of text and graphs in online learning has been noted to increase student engagement and understanding of course content (Al-Aghbari, Osman and Al Musawi 2021).

Video lectures include a variety of learner interactional components that are intended to enhance knowledge supplied by various types of media (Alraimi, Zo and Ciganek, 2015). The method in which media is employed in video lectures makes it one of the most important components, because it allows for most types of teaching and learner interaction, and online learning would not work without it (Lange and Costley, 2020). Similarly, research has also shown that visual media such as graphs, images, maps, slides, and text have been found to improve student attention and engagement during video lectures (Kizilcec, Bailenson and Gomez, 2015). Furthermore, when learners have to scan their eyes between texts and graphs, cognitive load is increased (Mayer, 2014), research suggests that the interplay between scanning text and graphs can significantly impact cognitive load (Raney, Campbell and Bovee 2014).

In addition, another study found that the graphical characteristics of printed text and the location and recognition of textual information can affect cognitive load (Zhong, et al., 2011). This is because when a graph is separated from the text that describes it, learners have to retain the textual information in their working memory while they search for the graph, thus increasing cognitive load and decreasing comprehension (Mayer, 2014). Low comprehension can lead to a lack of understanding within video lectures and a lack of understanding in video lectures might create a need to re-watch them. While no research has been found directly linking the arrangement of text and graphs to the need to re-watch videos, implications can be made from extant research. Mutlu-Bayraktar and Bayram (2018) showed greater levels of recall are evident when text is integrated with images as opposed to being physically separated to the point where students are forced to scan between the two sources. It has also been found that students tend to re-watch videos in which they have more difficulty processing information (Li, et al., 2015), thus the difficulty perceived through the increased load that leads to less recall when trying to process information from two physically separated sources may also lead to students re-watching the videos to make better sense of them.

2.4 Audiovisual Support

The use of audiovisuals in video lectures is becoming more frequently utilized in educational contexts. This is because audiovisual strengthens students' learning experiences by engaging a multitude of senses and promoting active learning (Nicolaou, Matsiola and Kalliris, 2019). Similarly, audiovisual media has been noted to facilitate the transmission of information between transmitter, and receiver and they enhance the learning process through representations (Rodriguez, 2007). Furthermore, Mayer (2014) argued that the utilization of both auditory and visual media to present instruction assists learners in comprehending lesson content.

Audiovisual support is the application of media within video lectures that enable learners to process information received through both visual and auditory modalities. While different studies have shown the importance of combining auditory and visual media to improve learning experiences, several studies have shown that learning can be negatively impacted by extraneous processing, which occurs when learners are distracted by the combination of multiple media sources (Lange and Costley, 2019; Leppink, et al., 2013). If the audiovisuals are presented in an engaging and non-distracting way, allowing learners to focus only on content that contributes to learning, it may reduce extraneous processing (Lange & Costley, 2020). However, studies also found that the separation of complementary media sources of information on the screen may lead to additional distractions, affecting the transmission of learning content negatively (Chen and Wu, 2015; Kizilcec, Bailenson and Gomez, 2015; Mayer, 2014). Similarly, the visual and auditory stimulation provided by audiovisual materials may act as a distraction for some students, leading to reduced focus and comprehension of the lecture material (Xiaojun, Zongkui and Zhongfeng, 2010). Considering this, it is important to understand the impact of audiovisual support in video lectures and how audiovisual support can affect the lecture rewatching behavior of the learners.

2.5 Professor's Speech Clarity

Clarity is essential to any good academic communication. To effectively engage students with video lectures in MOOCs, professors/instructors must prioritize clarity in their communication. This means that instructors should communicate their concepts simply, by utilizing suitable language, and avoid complexity or convoluted terminology (Guo, Kim and Rubin, 2014). Studies have indicated that unclear speech can disrupt the learning process, leading students to go back and replay the audio (Cunningham, Fägersten and Holmsten, 2010). Similarly, a study conducted by Leppink, et al. (2013) discovered that unclear explanations or instructions and unclear language contribute to an increase in extraneous load which in turn affects student learning processes.

In addition, when audio is unintelligible, it can result in learners not understanding what is being said by the instructor, which has been demonstrated to decrease comprehension (Pilarski, et al., 2008). Furthermore, the cognitive processes of learners have also been found to be influenced by the speed at which auditory media are delivered (Mayer, 2014). Some researchers found that instructors who speak fast and exhibit high levels of enthusiasm demonstrate higher levels of engagement (Guo, Kim and Rubin, 2014). Thus it can be concluded that while the physical presence of the instructor appears to benefit engagement levels, the way in which the information is delivered by instructors should be done in a way in which learners feel their time is being used in a productive and engaging manner. This falls in line with research that suggests students who do re-watch video lectures tend to focus on only the parts of the lecture relevant to their learning needs, skipping any unneeded information in the learning process (Kim, et al., 2014).

On the other hand, studies have shown that learners may be overloaded by fast-moving speech, which can thereby reduce comprehension (Koumi, 2013). Similarly, another study also stated that students can fail to gain critical information when learning from fast-paced media which can result in learning difficulty (Wildemuth, et al., 2003). When media is delivered at a fast pace it may result in low comprehension. The limited capacity theory explains the reduced level of comprehension when media is displayed at fast rates, stating that the brain's working memory has a limited ability to organize and process modality-specific information, and thus cognitive overload may take place when excessive information is delivered at a fast pace (Mayer, 2014). With different studies providing conflicting perspectives, it is needed to understand how professors' speech clarity in video lectures affects student learning and if the speech clarity for the professor/instructor can necessitate the need to rewatch video lectures. Additionally, although there is a lack of direct quantitative evidence that a relationship exists between instructor clarity and the need to re-watch videos, Fenton and Murphy (2023) did provide qualitative data that suggest that clarity is associated with the re-watching of instructor feedback in video lectures, as students reported higher levels of instructor understanding during the re-watch.

2.6 Text-Image Comprehensibility

In online learning, video lectures have become an essential instrument to deliver instruction to learners (Costley, et al., 2021). These video lectures often contain a combination of text and images to enhance comprehension and engage learners (Brame, 2016). On the other hand, studies have shown that the size of the text and image can affect the instructional delivery in video lectures. When the font size of a text is too small, it makes the content in instructional contexts illegible, thereby negatively affecting the comprehension levels of the learners (Amigud, et al., 2017; Sanchez and Goolsbee, 2010). Similarly, the size of images used in lecture content has also been noted to negatively affect student learning experiences (Al Ghamdi, et al., 2016; Molnar, 2017). When the learners find it hard to understand the content in the instruction due to its size, this may result in a situation where the students might have to rewatch these lectures to understand the lectures better.

Furthermore, research has also shown that when learners are presented with unclear instruction or instructional content that is difficult to understand, it may lead to extraneous load (Costley, et al., 2021; Leppink, et al., 2013; Schmeck, et al., 2015) making it difficult for the student to understand the lecture content. In order to gain a better understanding of the lecture content, students may engage in a variety of viewing strategies, some of which are rewatching the video lectures, splitting attention between text and images, and scanning text to find certain details required to obtain a better understanding of the content (Costley, et al., 2018; Kim, et al., 2014; Le, et al., 2010). Conversely, when the text and images used in video lectures are well-designed and easily comprehensible, students may be less likely to feel the need to re-watch the lectures. There has been no study found that fully explains the relationship between text and image comprehensibility in video lectures and its impact on lecture re-watching behavior by learners; however, given the evidence that unclear text and images lead to a more difficult learning experience (Al Ghamdi, et al., 2016; Molnar, 2017) coupled with the fact that students are more likely to re-watch video lectures that they perceive to be more difficult (Li, et al., 2015), it would make sense that students would be less likely to re-watch videos in which the text and images are easy to understand.

2.7 The Present Study

While previous studies have explored various aspects of student engagement in MOOCs, including the impact of instructional design, content delivery, and various strategies on student engagement, there is a noticeable gap in the literature concerning the direct relationship between specific instructional factors and the need to re-watch video lectures. This study aims to fill this gap by examining the unique instructional environments of the Open Cyber University (OCU) in South Korea, which offers a diverse range of courses with varying levels of instructional scaffolding and visual aids.

OCU's courses provide an excellent case for investigating these dynamics due to their distinctive instructional delivery methods, ranging from teacher-centered lectures with minimal visual aids to those incorporating extensive learner-controlled options and visual supports. Given that OCU primarily offers fully online classes, understanding which video lecture features prompt students to re-watch lectures can provide valuable insights for enhancing online learning experiences in South Korea.

This study seeks to delve into the specific instructional factors at OCU that influence students' decisions to re-watch video lectures. Understanding the video lecture features prompting students to re-watch video lectures is crucial for enhancing the effectiveness of student engagement with video lectures in MOOCs. For these reasons the present study has these four research questions:

- RQ1: How does the frequency of scanning between text and graphs influence the necessity for students to re-watch video lectures in MOOCs?
- RQ2: What is the impact of audio-visual support on the need for students to re-watch video lectures in MOOCs?
- RQ3: How does the ease of understanding textual and visual elements (text and images) affect the likelihood of students needing to re-watch video lectures in MOOCs?
- RQ4: To what extent does the clarity of the professor's speech contribute to students' decisions to re-watch video lectures in MOOCs?

2.8 Hypotheses

There is a positive relationship between the frequency of scanning eyes between text and graphs and the need to re-watch video lectures.

There is a significant impact of audiovisual distraction on re-watching video lectures.

There is a positive relationship between comprehensibility of text and images and the need to re-watch video lectures.

There is a significant relationship between the clarity of the professor's speech on the need to rewatch the video lectures.

3. Research Methodology

3.1 Research Design

In this study's context, survey participants were queried regarding their experiences with courses offered by the Open Cyber University (OCU), which provides approximately 400 different courses to a yearly enrolment of around 120,000 students (Han, 2012). While most OCU lectures feature an instructor appearing as a talking head in front of a plain backdrop without any supplementary visual aids, other OCU lectures incorporate visual aids alongside a range of learner-controlled options, allowing students to tailor their learning experiences to their preferences. From the standpoint of instructional delivery, the level of scaffolding provided varies from one class to another, with some instructors employing more comprehensive sequencing and fading of instruction (gradual reduction of instructional support as students gain mastery over the material being taught.) than others. It's worth noting that OCU primarily offers fully online classes, and it is essential to recognize that fully online classes in South Korea typically adhere to a conventional, teacher-centered approach (Lim, Kang and Park, 2016). This underscores the importance of improving the online learning experience in the form of video lectures for students in South Korea.

3.2 Procedures and Participants

The survey items were originally drafted in English and then translated into Korean. Subsequently, an expert in e-learning and the Open Cyber University (OCU) meticulously verified the translated content for accuracy. The OCU's ethics board, responsible for evaluating the research's ethical merit and its permissible use, granted approval for the survey. The survey link was subsequently included in an invitation extended to students, inviting their participation in the research. The survey was administered over a span of four weeks, following which the collected data was input into SPSS for subsequent analysis. A total of 1545 surveys were submitted, but 7 of these proved to be incomplete, resulting in 1538 usable surveys. An initial data examination focused on identifying outliers using Cook's, Mahalanobis, and Leverage values within the primary variables. In order to detect influential observations and potential outliers, Cook's Distance and Mahalanobis Distance were applied. Cook's Distance identifies points that significantly affect regression results (Cook, 1977), while Mahalanobis

Distance considers multivariate outliers by accounting for correlations between variables (Mahalanobis, 1936). The robustness of the dataset and model was ensured by these metrics As a result of this scrutiny, 9 cases were excluded, leaving a dataset of 1529 valid cases. The tables and figures presented in the forthcoming analysis are based on this refined dataset.

To ensure the validity and reliability of the data collected from the study sample, IBM SPSS Statistics (version 23) software was used to conduct reliability tests on the data by the researcher. Cronbach's alpha test was used to measure the internal consistency among the study variables. The analysis of the data found that the scale had a good internal consistency, with a Cronbach's alpha coefficient of 0.77. Cronbach's alpha is a measure of internal consistency, indicating how well the items on a scale are related to one another. A value of 0.77 suggests that the scale is reliable, meaning that the items are measuring the same construct consistently.

The study participants were requested to respond to the survey in relation to their use of videos while studying and learning at the OCU. The dataset included representations from 126 distinct classes, which could be categorized according to the OCU's classification as follows: lifestyle and health (29%), social science (26%), humanities (11%), business and management (8%), computers and information technology (8%), foreign language (7%), natural science (7%), and mathematics (4%). While each class did not have exactly the same type of teaching materials and number of videos, the OCU operates on a 15 week teaching cycle and most courses had 15 corresponding videos, with each video averaging approximately 20 to 30 minutes in length. The participant composition featured 872 females (56.8%) and 657 males (43.2%), with an average age of 23.7. The youngest participant was 18, while the oldest was 63. It's notable that this distribution of age, gender, and field of study aligns with the typical demographics found in research on e-learning environments in South Korea (Suh and Kim, 2013).

Table 1 describes four constructs reflecting specific features of the video lectures that were created based on the questions where Likert-type scale with the range from 0 to 10, with 0 being 'strongly disagree' and 10 being 'strongly agree'.

Table 1: Table showing the research questions and their constructs.

Questions	**Construct**
I had to scan my eyes between text and graphs	Arrangement of text and graphs
Audiovisuals are distracting (reversed)	Audiovisual support
The text and images are easy to understand	Text-image comprehensibility
When the professor speaks it's easy to understand him	Professor's speech clarity
Sometimes I had to rewatch part of the video	Video lecture rewatching behavior

4. Results and Findings

4.1 Overview of Main Findings

Multiple linear regression analysis at 95% confidence intervals was used for the data analysis. The dependent variable "Video lecture rewatching behavior" was regressed on predicting variables: arrangement of text and graphs, audiovisual support, text-image comprehensibility, and professor's speech clarity. The independent variables significantly predict "video lecture re-watching" as the analysis showed a good model fit $F_{(4, 1524)}$ = 42.928, $p < 0.01$, which indicates that the four factors under study have a significant impact on video lecture re-watching behavior. Variance Inflation Factor (VIF) and Tolerance values were examined. All predictors had VIF values below 3, which is well within the accepted threshold indicating no severe multicollinearity (Hair et al., 2010), and the tolerance values are consistently above 0.1, which further supports the absence of multicollinearity concerns. The Durbin-Watson statistic (1.793) confirmed the independence of residuals.

As shown in Figure 1, H1 evaluates whether there is a positive relationship between the arrangement of text and graphs and the need to re-watch video lectures. The results showed that the arrangement of text and graphs exhibited a significant positive relationship with the frequency of rewatching parts of the video (β = 0.34, t =

12.10, p < 0.01). Hence, H1 was supported. These results suggest that increased scanning between text and graphs is associated with a higher likelihood of needing to re-watch parts of the video, potentially indicating comprehension difficulties or cognitive overload during the learning process. H2 evaluates whether there is a significant impact of audiovisual support on re-watching video lectures, the relationship between the distraction caused by audiovisual support and the frequency of rewatching parts of the video was not statistically significant ($\beta = 0.05$, $p = 0.08$), suggesting that this variable has a weaker impact on the need to rewatch video lectures. Consequently, H2 was rejected. The results are presented in Table 2 and Figure 1.

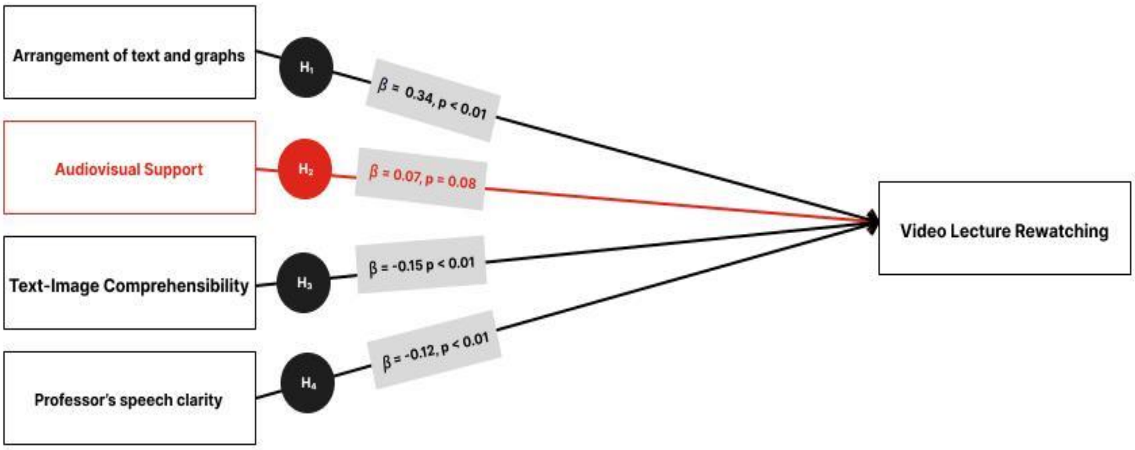

Figure 1: A diagram of multilinear regression analysis performed to evaluate the predictors and dependent variables

Another predictor variable, H3 evaluates if there is a significant positive impact on the comprehensibility of text and images with the need to re-watch video lectures. text-image comprehensibility demonstrated a statistically significant negative relationship with the frequency of rewatching parts of the video lectures ($\beta = -0.14$, $t = -3.39$, $p < 0.01$). Hence, H3 was supported. This shows that when learners perceive the texts and images in the video lectures as easy to understand, there is a decrease in the frequency of re-watching parts of the video lectures. The negative coefficient (-0.15) indicates that as the ease of understanding increases, the need to re-watch parts of the video lectures decreases. This relationship is statistically significant ($p < 0.01$), suggesting that it is unlikely to have occurred by chance.

Lastly, H4 assesses whether there is a significant impact on the clarity of the professor's speech on the need to re-watch the video lectures. The professor's speech clarity also showed a statistically significant negative association with the need to rewatch parts of the video lecture ($\beta = -0.12$, $t = -2.81$, $p < 0.01$). The results showed that when learners find it easy to understand the professor's speech in the video lectures, there is a decrease in the frequency of re-watching parts of the lectures. The negative coefficient (-0.12) suggests that clearer speech patterns contribute to a reduced need for re-watching video lectures. This relationship is statistically significant ($p < 0.01$), indicating its reliability. The results are presented in Table 2.

Table 2: Result of the multiple linear regression performed analysis on the dependent and dependent variables

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	4.932	.289		17.086	.000	4.366	5.498		
	The texts images and graphs are easy to understand	-.147	.043	-.141	-3.397	.001	-.233	-.062	.341	2.936

Coefficients										
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
	When the professor speaks it is easy to understand him	-.120	.043	-.117	-2.807	.005	-.204	-.036	.340	2.944
	audiovisuals are distracting (reversed)	.068	.039	.049	1.767	.077	-.008	.144	.783	1.278
	I had to scan my eyes between text and graphs	.339	.028	.323	12.102	.000	.284	.394	.830	1.205

a. Dependent Variable: Sometimes I had to rewatch parts of the video

4.2 Discussion

The findings from this study reveal key insights into the factors that significantly influence the need for students to re-watch video lectures in MOOCs, highlighting the multifaceted nature of student engagement with video lectures. Some of the factors that were found to have a significant impact on the need to re-watch video lectures include arrangement of text and graphs, text-image comprehensibility, and the professor's speech clarity. These results suggest that arranging text and graphs to minimize eye scanning, ensure clarity of the professor's speech, and enhance text-image comprehensibility contributes to effective video lecture-watching strategies.

One of the most noteworthy findings is the positive relationship between the arrangement of text and graphs (how frequently students have to eye scan between text and graphs) and the need to rewatch video lectures. This finding suggests that frequent eye-scanning between text and graphs may contribute to cognitive overload, compelling students to re-watch parts of the video. This finding aligns with the cognitive load theory and two of its effects: redundancy effect occurring when the learners are presented with the same information in more than one modality (Sweller, 2024) and the split attention effect that promotes replacing multiple sources of information with one (van Merrienboer and Ayres, 2005). Students' working memory resources are depleted as they scan both the text and graphs and might go back and forth in order to understand the content. Previous studies support this by showing that students expend additional cognitive resources when integrating fragmented information (Mutlu-Bayraktar & Bayram, 2018). As a result of this they might therefore need to re-watch part of the video lectures if they have poor comprehension of the text and graphs used in this type of instructional content. Recent studies corroborate this finding. For instance, Raney, Campbell, and Bovee (2014) demonstrated that integrating text and graphical elements in instructional videos significantly reduces cognitive load and enhances information retention. The results suggest that designers of MOOC content should prioritize the seamless integration of textual and graphical elements to mitigate cognitive strain.

When creating and designing lecture content, instructors should employ strategies promoting optimizing the visual representation of the content to not create additional cognitive load that could impede learning. In addition, educators should strive to design instructional materials that minimize the need for extensive eye-scanning, perhaps by integrating text and graphical information more seamlessly or using annotations to guide students' attention effectively. This finding also suggests that designers of educational content should consider the principles of multimedia learning when creating instructional videos. For instance, placing text next to relevant graphics, rather than separately, can help reduce the cognitive effort required to integrate the two types of information, ultimately leading to greater recall (Mutlu-Bayraktar and Bayram, 2018). Moreover, interactive elements, such as clickable diagrams that provide additional information when needed, could help manage cognitive load by allowing students to access information at their own pace.

In addition, another finding shows that there is a significant impact of the comprehensibility of text and images on the need to re-watch video lectures. According to our findings, when learners perceive texts and images as easy to understand, they are less likely to re-watch parts of the video lectures. This finding aligns with cognitive load theory's assertion that reducing extraneous load—such as poor visual clarity—can improve learning efficiency (Mayer, 2014).This result underscores the importance of clear and comprehensible visual aids in instructional materials as the visual clarity of instructional content emerged as a critical determinant influencing students' propensity to re-watch video lectures. While Mayer and Moreno (2003) argued that excessive audiovisual elements can lead to cognitive overload, the present findings suggest that audiovisual distractions may not play a pivotal role in students' re-watching behavior. However, another study supports this finding by indicating that poor visual clarity can hinder students' comprehension, necessitating multiple viewings to grasp the content fully (Molnar, 2017). Based on this finding, it has been identified that poor visual clarity can impede students' understanding of video lectures, necessitating the need to re-watch video lectures to achieve a deeper grasp of the content. Low comprehension resulting from poor visual clarity can perpetuate a cycle wherein students feel compelled to revisit lectures in pursuit of deeper understanding. This finding is in line with another research that noted that students struggle to comprehend images and understand text used in media when the media used in instruction is very fast with low visual clarity thereby resulting in low comprehension as this prevents students from obtaining critical information from the media (Wildemuth, et al., 2003). This aligns with the cognitive load theory, suggesting that reducing extraneous cognitive load by enhancing visual clarity can improve learning efficiency (Mayer, 2014). Moreover, the implications of this finding extend to the design of instructional materials in general. Effective visual aids should not only be clear but also well-integrated with the spoken and written components of the lecture. This integration can help prevent the cognitive overload that occurs when students have to mentally integrate disparate sources of information.

Furthermore, creating content aimed at enhancing the visual clarity of instructional materials holds promise for alleviating the burden of repetitive review and empowering students to engage more meaningfully with the course content (Koh and Ahn, 2023). Conversely, when presented with clear and comprehensible text and images, students exhibit a reduced demand for revisiting video lectures. This highlights the pivotal role of clear and comprehensible visual aids in instructional materials and video lectures within MOOCs. Designing content with a judicious balance between complexity and clarity becomes imperative in optimizing student comprehension and minimizing the need for redundant review of course materials.

Furthermore, the current study found that learners' re-watching behavior of video lectures was impacted by the professor's speech clarity. Clear and comprehensible instructor speech reduces the likelihood that students will need to re-watch parts of the lecture. This finding is in line with previous research showing that video lectures with fast- speakers might lead to students becoming confused and overwhelmed, which in turn provokes them to replay parts of video lectures (Guo, Kim and Rubin, 2014). Similarly, another study showed that when the audio is audible and intelligible, students are unlikely to replay the lecture (Cunningham, Fägersten and Holmsten, 2010). Furthermore, it has also been demonstrated that problems with auditory intelligibility can lead to a reduction in comprehension levels because students may be unable to understand what is being said. And when learners find it hard to understand some part of the lecture due to its auditory intelligibility, they may need to rewatch the lecture to understand it better. Results that are more directly tied to the clarity of the instructor specifically support the findings of this study, as it has been found that students perceive more instructor clarity when re-watching feedback portions of video lectures (Fenton, and Murphy, 2023). In practical terms, this means that instructors should aim for a moderate speaking pace and use clear, concise language. Additionally, they can benefit from using subtitles and transcripts, which have been shown to aid comprehension, especially for non-native speakers (Pilarski, et al., 2008). Providing these additional resources can help ensure that all students, regardless of their language proficiency, can understand the lecture content without needing to re-watch it multiple times. These findings, along with the results of the current study, suggest that in the context of MOOCs, learners may not need to re-watch video lectures if once the speech voice of the instructor in the video lectures is clear, and they can understand the speech easily.

Lastly, results from the study also noted that there is little to no significant impact of audiovisual support on the re-watching behavior of the learners on the MOOC platform. The study found that audiovisual support did not significantly impact the need to re-watch video lectures. While cognitive load theory suggests that excessive use of audiovisuals can lead to cognitive overload (Mayer and Moreno, 2003), another study by Lackmann, et al., 2021, found that audiovisual support in video lectures led to better engagement, impact, performance, and learning outcomes, indicating their effectiveness in fostering deeper understanding. The current study's findings imply that the presence or absence of audiovisual support alone may not be a decisive factor in students' re-watching behavior. This discrepancy could be due to the specific context of OCU's courses, where instructional styles and content delivery methods vary. It is also possible that the type and quality of audiovisual materials play a role in their effectiveness. For example, high-quality visuals that are directly relevant to the content being discussed may be more beneficial than generic or distracting visuals. Additionally, the timing and pacing of audiovisual elements can influence their effectiveness. Well-timed visuals that reinforce the spoken content can help enhance understanding, whereas poorly timed or irrelevant visuals can cause confusion and distraction which might result in a situation where students need to rewatch the video lecture for better understanding.

Other studies confirm that the video lecture style "talking head" in which the instructor speaks directly into the camera while using little to no audiovisuals is the most effective method of video lecture production style (Ilioudi, Giannakos and Chorianopoulos, 2013). This underscores the significance of minimizing audiovisuals in video lectures as the enormous use of audiovisuals can cause cognitive overload and unnecessary distraction which can lead to the learners re-watching the video lectures. Educators and content creators for MOOCs should consider designing courses with a focus on minimizing audiovisual distractions to foster sustained engagement.

4.3 Limitations

While this study sheds light on critical aspects of video lecture engagement and the re-watching behavior of video lectures in MOOCs, certain limitations must be acknowledged. First of all, the survey methodology relies on self-reported data of a limited range of video features. While the internal consistency of the instruments used in this study was confirmed to be reliable, the results of this study could be further validated in future research through experimental conditions. Additionally, the study predominantly examines quantitative aspects, leaving room for further qualitative exploration of student engagement with video lectures and the need to re-watch video lectures in MOOCs. Further research encompassing diverse learning platforms and a mixed-methods approach could provide a more comprehensive understanding of student engagement and re-watching behavior in MOOCs. Finally, in order to conclude whether the phenomenon of re-watching video lectures could serve as a barometer of the instructional efficacy of video lectures, it is important to investigate its impact on the learning outcomes. Furthermore, future research could consider some of the factors mentioned in the present study on different populations to see if the results are generalizable.

5. Conclusion

In conclusion, this research contributes significant insights into the multifaceted nature of student engagement with video lectures in MOOCs and the underlying factors driving the need to re-watch video lectures. Based on the findings of this study the following instructional guidelines are recommended: 1) Present content that minimizes the need for extensive eye-scanning by positioning complementary textual/graphical sources in close proximity to each other. 2) Integrate text and graphical information more seamlessly or use annotations to guide students' attention effectively. 3) Clearly integrate visuals with spoken and written components of the lecture by ensuring that they complement each other rather than provide redundant or distracting information. 4) Deliver instruction in a clear and concise language that only applies to the specific learning tasks to be learned at a given time. Following this framework should help learners avoid the need to re-watch videos.

Our findings underscored the pivotal role of specific video attributes in shaping students' decisions to re-watch video lectures. Notably, the arrangement of text and graphs, the professor's speech clarity, and text-image comprehensibility emerged as significant determinants influencing re-watching behavior. The associations between text-image comprehensibility and professor's speech clarity and the need to re-watch parts of video lectures elucidate the profound impact of clear instructional delivery on student comprehension and retention. Furthermore, the positive relationship observed between the arrangement of text and graphs and the frequency of re-watching underscores the potential challenges associated with cognitive processing and information assimilation in MOOC environments. However, the absence of a statistically significant relationship between audiovisual support and the need to re-watch video lectures suggests a nuanced interplay between various video characteristics and student engagement patterns. Overall, these findings provide valuable guidance for educators, content creators, and instructional designers seeking to optimize student engagement and learning outcomes in MOOCs.

Acknowledgements

This work/article is an output of a research project implemented as part of the Basic Research Program at the National Research University Higher School of Economics (HSE University).

AI statement: We, the authors, hereby declare that the work presented in this academic paper is the result of our collective research and efforts. We acknowledge the use of an AI (Quillbot AI and WordTune) to assist in paraphrasing sections of the text. The AI tool was employed to enhance the clarity and coherence of the writing, while all intellectual contributions and final decisions regarding the content were made by us. We confirm that the use of AI was conducted ethically and transparently, in accordance with the guidelines and standards of academic integrity. All sources and references have been appropriately cited, and the originality of the work has been maintained.

Ethics statement: The OCU's ethics board, responsible for evaluating the research's ethical merit and its permissible use, granted approval for the survey.

References

Al-Aghbari, M.S., Osman, M.E. and Al Musawi, A.S., 2021. Contextualizing the Global Standards for Designing Online Courses: A Design-Based Research Approach for Developing Small Private Open Courses. International Journal of Educational Methodology, 7(1), pp.1-13.

Abbakumov, D., Kravchenko, D., Kuskin, W. and Urban, A., 2020. How rewatching video lectures impacts students' performance in assessments in MOOCs. https://doi.org/10.35542/osf.io/vcd5g

Afıfy, M.K., 2020. Effect of interactive video length within e-learning environments on cognitive load, cognitive achievement and retention of learning. Turkish Online Journal of Distance Education, 21(4), pp.68-89.

Albelbisi, N.A., Al-Adwan, A.S. and Habibi, A., 2023. A qualitative analysis of the factors influencing the adoption of mooc in higher education. *Turkish Online Journal of Distance Education*, 24(2), pp.217-231. https://doi.org/10.17718/tojde.973956.

Al Ghamdi, E., Yunus, F., Da'Ar, O., El-Metwally, A., Khalifa, M., Aldossari, B. and Househ, M., 2016. The effect of screen size on mobile phone user comprehension of health information and application structure: an experimental approach. *Journal of medical systems*, 40(1), pp.1-8. https://doi.org/10.1007/s10916-015-0381-5.

Akcapinar, G. and Bayazit, A., 2018. Investigating video viewing behaviors of students with different learning approaches using video analytics. *Turkish Online Journal of Distance Education*, *19*(4), pp.116-125. https://doi.org/10.17718/tojde.471907.

Alraimi, K.M., Zo, H. and Ciganek, A.P., 2015. Understanding the MOOCs continuance: The role of openness and reputation. *Computers & Education*, *80*, pp.28-38. https://doi.org/10.1016/j.compedu.2014.08.006.

Amigud, A., Arnedo-Moreno, J., Daradoumis, T. and Guerrero-Roldan, A.E., 2017. Using learning analytics for preserving academic integrity. *International Review of Research in Open and Distributed Learning*, *18*(5), pp.192-210. https://doi.org/10.19173/irrodl.v18i5.3103.

Aparicio, M., Oliveira, T., Bacao, F. and Painho, M., 2019. Gamification: A key determinant of massive open online course (MOOC) success. *Information & Management*, *56*(1), pp.39-54. https://doi.org/10.1016/j.im.2018.06.003.

Atapattu, T. and Falkner, K., 2018. Impact of lecturer's discourse for students' video engagement: Video learning analytics case study of MOOCs. *Journal of Learning Analytics*, *5*(3), pp.182-197. https://doi.org/10.18608/jla.2018.53.12.

Banoor, R.Y., Issack, S.M. and Frank, R., 2019. The Use of Learning Analytics to Improve Online Learning Outcomes: A Systematic Literature Review.

Brame, C.J., 2016. Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education*. https://doi.org/10.1187/cbe.16-03-0125.

Bonafini, F.C., 2017. The effects of participants' engagement with videos and forums in a MOOC for teachers' professional development. *Open Praxis*, *9*(4), pp.433-447. https://doi.org/10.5944/openpraxis.9.4.637.

Burd, E.L., Smith, S.P. and Reisman, S., 2015. Exploring business models for MOOCs in higher education. *Innovative Higher Education*, *40*, pp.37-49. https://doi.org/10.1007/s10755-014-9297-0.

Chen, C.M. and Wu, C.H., 2015. Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, *80*, pp.108-121. https://doi.org/10.1016/j.compedu.2014.08.015.

Cojocariu, V.M., Lazar, I., Nedeff, V. and Lazar, G., 2014. SWOT analysis of e-learning educational services from the perspective of their beneficiaries. *Procedia-Social and Behavioral Sciences*, *116*, pp.1999-2003. https://doi.org/10.1016/j.sbspro.2014.01.510.

Cook, R.D., 1977. Detection of influential observation in linear regression. Technometrics, 19(1), pp.15-18.

Costley, J., Hughes, C. and Lange, C., 2017. The effects of instructional design on student engagement with video lectures at cyber universities. Journal of Information Technology Education: Research, 16(1) .https://doi.org/10.28945/3728.

Costley, J., Fanguy, M., Baldwin, M., Lange, C., & Han, S., 2018. The role of motivation in the use of lecture behaviours in the online classroom. *Journal of Information Technology Education Research*, *17*(1), pp.471–484. https://doi.org/10.28945/4131.

Costley, J., Fanguy, M., Lange, C. and Baldwin, M., 2021. The effects of video lecture viewing strategies on cognitive load. *Journal of Computing in Higher Education*, *33*, pp.19-38. https://doi.org/10.1007/s12528-020-09254-y.

Cunningham, U.M., Fägersten, K.B. and Holmsten, E., 2010. Can you hear me, Hanoi? Compensatory mechanisms employed in synchronous net-based English language learning. *The International Review of Research in Open and Distributed Learning*, *11*(1), pp.161-177. https://doi.org/10.19173/irrodl.v11i1.774.

Dhawan, S., 2020. Online learning: A panacea in the time of COVID-19 crisis. *Journal of educational technology systems*, *49*(1), pp.5-22. https://doi.org/10.1177/0047239520934018.

Egloffstein, M. and Ifenthaler, D., 2017. Employee perspectives on MOOCs for workplace learning. *TechTrends*, *61*(1), pp.65-70. https://doi.org/10.1007/s11528-016-0127-3.

Fenton, A., & Murphy, P., 2023. Student perceptions of video feedback: Preliminary Findings. ASCILITE Publications, 398-403.

Guo, P.J., Kim, J. and Rubin, R., 2014. March. How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 41-50). https://doi.org/10.1145/2556325.2566239.

Hair, J., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L., 2010. Multivariate Data Analysis. 6th (ed.) Upper Saddle River NJ.

Han, S.L., 2012. A study on the development of a course evaluation tool for Cyber University Consortium in Korea. In *International Conference on Human Computer Interaction* (pp. 331-337). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35270-6_46.

Hsin, W.J. and Cigas, J., 2013. Short videos improve student learning in online education. Journal of Computing Sciences in Colleges, 28(5), pp.253-259.

Hussain, M., Zhu, W., Zhang, W. and Abidi, S.M.R., 2018. Student engagement predictions in an e-learning system and their impact on student course assessment scores. *Computational Intelligence & Neuroscience*. https://doi.org/10.1155/2018/6347186.

Hwang, G.J., Lai, C.L. and Wang, S.Y., 2015. Seamless flipped learning: a mobile technology-enhanced flipped classroom with effective learning strategies. *Journal of computers in education*, *2*, pp.449-473. https://doi.org/10.1007/s40692-015-0043-0.

Ilioudi, C., Giannakos, M.N. and Chorianopoulos, K., 2013, April. Investigating differences among the commonly used video lecture styles. In *Proceedings of the Workshop on Analytics on Video-based Learning, WAVe* (Vol. 13, pp. 21-27).

Jin, L. and Shang, D., 2022. What drives art education massive open online courses continuance intention? Exploring determinants from value–software–hardware–design perspective. *Interactive Learning Environments*, pp.1-17. https://doi.org/10.1080/10494820.2022.2153146.

June, S., Yaacob, A. and Kheng, Y.K., 2014. Assessing the use of YouTube videos and interactive activities as a critical thinking stimulator for tertiary students: An action research. International Education Studies, 7(8), pp.56-67. https://doi.org/10.5539/ies.v7n8p56.

Kim, J., Guo, P.J., Seaton, D.T., Mitros, P., Gajos, K.Z. and Miller, R.C., 2014. Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference* 31-40. https://doi.org/10.1145/2556325.2566237.

Kizilcec, R.F., Bailenson, J.N. and Gomez, C.J., 2015. The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology*, *107*(3), p.724. https://doi.org/10.1037/edu0000013.

Koh, T. and Ahn, J., 2023. The effects of student-engaged video lectures on motivation for sustainable flipped learning. Sustainability, 15(5), p.4617.

Koumi, J., 2013. Pedagogic design guidelines for multimedia materials: A call for collaboration between practitioners and researchers. Journal of Visual Literacy, 32(2), pp.85-114.

Kuznekoff, J.H., 2020. Online video lectures: the relationship between student viewing behaviors, learning, and engagement. AURCO Journal, 26, pp.33-55.

Lackmann, S., Léger, P.M., Charland, P., Aubé, C. and Talbot, J., 2021. The influence of video format on engagement and performance in online learning. Brain Sciences, 11(2), p.128.

Lange, C. and Costley, J., 2019. The negative impact of media diversity on self-regulated learning strategies and cognitive load. *Issues in Educational Research*, *29*(1), pp.158-179. http://www.iier.org.au/iier29/lange.pdf

Lange, C. and Costley, J., 2020. Improving online video lectures: learning challenges created by media. *International Journal of Educational Technology in Higher Education*, *17*, pp.1-18. https://doi.org/10.1186/s41239-020-00190-6.

Le, A., Joordens, S., Chrysostomou, S. and Grinnell, R., 2010. Online lecture accessibility and its influence on performance in skills-based courses. *Computers & Education*, *55*(1), pp.313-319. https://doi.org/10.1016/j.compedu.2010.01.017.

Leppink, J., Paas, F., Van der Vleuten, C.P., Van Gog, T. and Van Merriënboer, J.J., 2013. Development of an instrument for measuring different types of cognitive load. *Behavior research methods*, *45*, pp.1058-1072. https://doi.org/10.3758/s13428-013-0334-1.

Li, N., Kidzinski, L., Jermann, P., & Dillenbourg, P., 2015. How do in-video interactions reflect perceived video difficulty?. Proceedings of the European MOOCs stakeholder summit 2015, 112-121.

Lim, K., Kang, M. and Park, S.Y., 2016. Structural relationships of environments, individuals, and learning outcomes in Korean online university settings. *International Review of Research in Open and Distributed Learning*, *17*(4), pp.315-330. https://doi.org/10.19173/irrodl.v17i4.2500.

Mahalanobis, P.C., 2018. On the generalized distance in statistics. Sankhyā: The Indian Journal of Statistics, Series A (2008-), 80, pp.S1-S7.

Maloney, S., Axelsen, M., Stone, C., Galligan, L., Redmond, P., Brown, A., Turner, J. and Lawrence, J., 2023. Defining and exploring online engagement fatigue in a university context. Computers and Education Open, 4, p.100139.

Martin, L., Mills, C., D'Mello, S.K. and Risko, E.F., 2018. Re-watching lectures as a study strategy and its effect on mind wandering. Experimental Psychology.65(5), pp.297-305.

Mayer, R.E., 2014. Multimedia instruction. *Handbook of research on educational communications and technology*, pp.385-399. https://doi.org/10.1007/978-1-4614-3185-5_31.

Mayer, R.E. and Moreno, R., 2003. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, *38*(1), pp.43-52. https://doi.org/10.1207/s15326985ep3801_6.

Molnar, A., 2017. Content type and perceived multimedia quality in mobile learning. *Multimedia Tools and Applications*, *76*, pp.21613-21627.

Mutlu-Bayraktar, D., & Bayram, S. (2018). Evaluation of situations causing split of attention in multimedia learning environments via eye-tracking method. In Digital Multimedia: Concepts, Methodologies, Tools, and Applications (pp. 348-372). IGI Global.

Nicolaou, C., Matsiola, M. and Kalliris, G., 2019. Technology-enhanced learning and teaching methodologies through audiovisual media. *Education Sciences*, *9*(3), p.196. https://doi.org/10.3390/educsci9030196.

Palmer, S., Chu, Y. and Persky, A.M., 2019. Comparison of rewatching class recordings versus retrieval practice as post-lecture learning strategies. American Journal of Pharmaceutical Education, 83(9), p.7217.

Patel, B., Yook, G., Mislan, S. and Persky, A.M., 2019. Exploring the consequences on memory of students who know they have access to recorded lectures. American journal of pharmaceutical education, 83(5), p.6958.

Pilarski, P.P., Johnstone, A. D., Pettepher, C.C. and Osheroff, N., 2008. From music to macromolecules: Using rich media/podcast lecture recordings to enhance the preclinical educational experience. *Medical teacher*, *30*(6), pp.630-632. https://doi.org/10.1080/01421590802144302.

Pozón-López, I., Kalinic, Z., Higueras-Castillo, E. and Liébana-Cabanillas, F., 2020. A multi-analytical approach to modeling of customer satisfaction and intention to use in Massive Open Online Courses (MOOC). *Interactive Learning Environments*, *28*(8), pp.1003-1021. https://doi.org/10.1080/10494820.2019.1636074.

Rackaway, C., 2012. Video killed the textbook star?: Use of multimedia supplements to enhance student learning. *Journal of Political Science Education*, *8*(2), pp.189-200. https://doi.org/10.1080/15512169.2012.667684.

Raney, G.E., Campbell, S.J. and Bovee, J.C., 2014. Using eye movements to evaluate the cognitive processes involved in text comprehension. J*ournal of Visualized Experiments*, 83. https://doi.org/10.3791/50780.

Rodriguez, S., 2007. *A practical model for multimedia development*. *Society for Information Technology & Teacher Education International Conference.* Texas, USA. March 6 2027. North Carolina: Association for the Advancement of Computing in Education.

Roehling, P.V., 2018. *Flipping the college classroom: An evidence-based guide*. Springer. https://doi.org/10.1007/978-3-319-69392-7_7.

Sanchez, C.A. and Goolsbee, J.Z., 2010. Character size and reading to remember from small displays. *Computers & Education*, *55*(3), pp.1056-1062. https://doi.org/10.1016/j.compedu.2010.05.001.

Sinha, T., Jermann, P., Li, N. and Dillenbourg, P., 2014. Your click decides your fate: Inferring information processing and attrition behavior from mooc video clickstream interactions. In *Proceedings of the EMNLP 2014 Workshop on Analysis of Large Scale Social Interaction in MOOCs.* pp. 3-14. https://doi.org/10.3115/v1/w14-4102.

Scagnoli, N.I., McKinney, A. and Moore-Reynen, J., 2015. Video lectures in eLearning. In *Handbook of research on Innovative Technology Integration in higher education.* pp. 115-134. https://doi.org/10.4018/978-1-4666-8170-5.ch006.

Seaton, D.T., Bergner, Y., Chuang, I., Mitros, P. and Pritchard, D.E., 2014. Who does what in a massive open online course?. *Communications of the ACM*, *57*(4), pp.58-65. https://doi.org/10.1145/2500876.

Schmeck, A., Opfermann, M., Van Gog, T., Paas, F. and Leutner, D., 2015. Measuring cognitive load with subjective rating scales during problem solving: differences between immediate and delayed ratings. *Instructional Science*, *43*, pp.93-114. https://doi.org/10.1007/s11251-014-9328-3.

Shah, D., 2018. By the numbers: MOOCs in 2018 central's MOOC report. [Online]. Available at https://www.classcentral.com/report/MOOCs-2015-stats/ [Accessed January 1, 2024].

Shao, Z. and Chen, K., 2021. Understanding individuals' engagement and continuance intention of MOOCs: the effect of interactivity and the role of gender. *Internet Research*, *31*(4), pp.1262-1289. https://doi.org/10.1108/intr-10-2019-0416.

Stockwell, B.R., Stockwell, M.S., Cennamo, M. and Jiang, E., 2015. Blended learning improves science education. *Cell*, *162*(5), pp.933-936. https://doi.org/10.1016/j.cell.2015.08.009.

Sweller, J., 2022. The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. Educational Psychology Review, 34(4), 2229-2241.

Sweller, J., 2024. Cognitive load theory and individual differences. *Learning and Individual Differences*, *110*, p.102423. https://doi.org/10.1016/j.lindif.2024.102423.

Suh, S. and Kim, S., 2013. Study on policy for an entrance quota of cyber universities. *Korea Educational Information and Research Service.*

Triay, J., Sancho-Vinuesa, T., Minguillón, J. and Daza, V., 2016. Analyzing non-linear video usage in an introductory x-MOOC about basic linear algebra.

Tzeng, J.W., Lee, C.A., Huang, N.F., Huang, H.H. and Lai, C.F., 2022. MOOC evaluation system based on deep learning. *International Review of Research in Open and Distributed Learning*, *23*(1), pp.21-40. https://doi.org/10.19173/irrodl.v22i4.5417.

Van Merrienboer, J.V. and Ayres, P., 2005. Research on Cognitive Load Theory and Its Design Implications for E-learning: Introduction to the special issue. *Educational Technology Research and Development*, 53, pp. 5–13. https://doi.org/10.1007/bf02504793.

Wildemuth, B.M., Marchionini, G., Yang, M., Geisler, G., Wilkens, T., Hughes, A. and Gruss, R., 2003. How fast is too fast? Evaluating fast forward surrogates for digital video. In *2003 Joint Conference on Digital Libraries, 2003. Proceedings.* pp. 221-230 . IEEE. https://doi.org/10.1109/jcdl.2003.1204866.

Xiaojun, L., Zongkui, Z. and Zhongfeng, X., 2010. The Effect of Audio-visual Teaching on Learning Burnout. In *2010 International Conference on E-Business and E-Government.* pp. 1669-1671 . IEEE. https://doi.org/10.1109/icee.2010.422.

Zhang, J., Gao, M. and Zhang, J., 2021. The learning behaviours of dropouts in MOOCs: A collective attention network perspective. *Computers & education*, *167*, p.104189. https://doi.org/10.1016/j.compedu.2021.104189.

Zhong, N., Li, M., Wu, Y. and Lu, S., 2011, May. The impact of different forms of statistical information on reading efficiency, effect, and mental workload: An eye-tracking study. In *The 2011 IEEE/ICME International Conference on Complex Medical Engineering.* pp. 97-102. IEEE. https://doi.org/10.1109/iccme.2011.5876712.

Foresight Study on the Influence of ChatGPT In Peruvian Universities Towards 2033

Mariel Gabriela Muñiz-Apaza[1], Julio César Acosta-Prado[2] and Carlos Guillermo Hernández-Cenzano[2,3]

[1]Graduate School, Pontifical Catholic University of Peru, Lima-Peru
[2]Department of Engineering, Pontifical Catholic University of Peru, Lima-Peru
[3]Systemic Planning Research Group, Pontifical Catholic University of Peru, Lima-Peru

m.muniz@pucp.edu.pe
jcacosta@pucp.edu.pe
Carlos.hernandez@pucp.edu.pe

https://doi.org/10.34190/ejel.23.2.3837

Abstract: This article investigates how the integration of ChatGPT could transform Peruvian universities by 2033. Twenty drivers were identified, including digital competencies, critical thinking, academic performance, and motivation. Using a foresight approach, possible future scenarios were generated. The methodology consists of four stages: (1) Exploration of the system to identify key drivers; (2) Validation of the information by applying the Delphi method in real time to confirm these drivers; (3) Construction of scenarios using Schwartz axes and structural analysis; and (4) Validation of scenarios using the Probability, Desirability and Governance (PDG) method. The analysis revealed that the most important and uncertain drivers are related to 'training in critical thinking, feedback through intelligent tutoring, and the development of technological infrastructure,' located in Schwartz's Quadrant III. These drivers are crucial to building four scenarios, including the target scenario dubbed "Overcoming Challenges with Innovation and Technology".

Keywords: Foresight study, Future scenarios, University, ChatGPT

1. Introduction

Technological advances, particularly in the field of artificial intelligence (AI), have opened new pathways for progress across all sectors and have revolutionized global education due to their rapid integration and the widespread attention they have received (Brunner and Tedesco, 2003). This research focuses on the study of generative artificial intelligence such as ChatGPT, which was selected for this research because among the many artificial intelligence algorithms this Open AI application is revolutionizing the academic field, generating challenges and opportunities.

ChatGPT has been applied in fields such as software engineering, health sciences, social sciences, agriculture, and marketing, among other sectors. However, in the education sector, it has offered more personalized and interactive learning (Sallam, 2023). Its integration into these fields not only promotes innovation but also fosters creative opportunities and solutions. It enables the creation and design of more personalized learning experiences for students and educators (Ali, et al. 2024). Its potential use in education lies in the personalization of learning, the automation of personalized tutorials, and the optimization of pedagogical feedback. However, its implementation in higher education poses several challenges, including ethical concerns, changes in teaching-learning dynamics, and the need for teacher training in these new technologies (Kuleto, et al. 2021).

In the Peruvian context, integrating artificial intelligence (AI) in Peruvian universities offers both opportunities and obstacles, and there has been a growing interest in digital transformation in recent times. However, institutions face disparities in relation to technological infrastructure and especially in internet access, limiting the benefits of technological resources, especially in rural areas (Luna, 2023).

This implies that there must be an improvement in digital inclusion to close the educational gap and improve the quality of education in the country.

The SITEAL and UNESCO report (2023) shows that 43% of Peruvian public universities lack adequate access to digital platforms, which creates a disparity between urban and rural institutions (Alfaro and Del Río, 2018). Other barriers in technological adoption, identified by Vidaurre et al. (2024), include the resistance to change among some teachers in using technological resources and lack of literacy or training in Artificial intelligence.

Cite this article: Muñiz-Apaza, M.G, Acosta-Prado, J.C and Hernández-Cenzano, C.G. 2025. "Foresight Study on the Influence of ChatGPT In Peruvian Universities Towards 2033", *Electronic Journal of e-Learning*, 23(2), pp 80-98, https://doi.org/10.34190/ejel.23.2.3837

AI has great potential to improve learning. However, regulatory frameworks must be developed to promote its ethical use. These frameworks should ensure a balanced approach, combining adoption with pedagogical strategies that foster critical thinking and ethical AI usage (Guadalupe et al., 2023).

On the other hand, the COVID-19 pandemic has highlighted the technological gaps in Peru. Areas with lack of internet access in many cases do not have adequate devices for virtual education, weakening equitable and quality access. The Peruvian education system faces challenges in terms of National education policies must ensure inclusive, equitable, and high-quality education (SITEAL and UNESCO, 2023).

Academic staff and students should be trained to prioritize critical thinking and ethical use. Studies highlight critical thinking as a skill rather than a disposition, showing greater effectiveness with instrumental strategies that align teaching competencies with the needs of students, providing didactic tools for future teaching performance (Andreucci-Annunziata, et al. 2023). To support the development of critical thinking, it is important to have different versions or points of view, allowing for multiple perspectives. In this sense, tools like ChatGPT provide valuable support by facilitating different approaches to topics, enhancing students' cognitive skills.

Given this panorama, this study seeks to investigate how the integration of ChatGPT can be transformed into Peruvian universities by 2033. To this end, possible future scenarios are generated, through a prospective approach, identifying factors of change that guide strategic decisions in the educational field. Through the Delphi method in real time and the construction of scenarios based on uncertainty axes Based on this, the necessary conditions for an effective implementation of AI in Peruvian higher education are analyzed.

The methodology involves four main stages: (1) Exploration of the system to identify the drivers; (2) Validation of information through the application of the Delphi method in real time (Astigarraga, 2003); (3) Construction of scenarios using Schwartz's axes and structural analysis, referring to ten years of study to systematically analyze the importance and uncertainty of the drivers, generating various scenarios (Schwartz, 1991). These axes are grouped into two main categories: "Promotion of critical thinking" and "Advanced technological infrastructure", essential for academic and professional development. (4) Validation of scenarios through the Probability, Desirability and Governance method (PDG). Finally, backcasting was used to formulate strategies and target scenarios, providing a solid framework for long-term strategic planning and policy development towards 2033 (Höjer and Mattsson, 2000).

2. Literature Review

2.1 Scenarios

In strategic planning, scenarios are valuable tools that allow organizations to explore and understand possible futures by generating scenarios to anticipate opportunities and identify risks. Scenarios are not predictions, but narratives that describe different possible futures based on trends. Scenarios are crucial for imagining futures and understanding individual and organizational change. This exploration helps to envision desirable futures and provides a framework for understanding change at the organizational level. In addition, scenarios are integrated with other methods, such as strategic forecasting, to provide a deeper understanding of the issue at hand (Inayatullah, 2008).

The value of scenarios lies in their ability to outline future outcomes that provide clarity and improve decision-making, transforming them into a tool for problem-solving (Wack, 1985). Integrating learning and planning from the outset is essential, with methods that consider the different perspectives of stakeholders and their evolution over time. Scenario generation increases awareness of upcoming changes, early warning signs, and new opportunities, while enriching participants' understanding of the issues (Graham, 2009). From an educational point of view, scenarios play a key role in the formulation of strategies for the adoption of new technologies, such as artificial intelligence, ensuring a more informed and adaptive approach in their implementation.Generative Artificial Intelligence (GAI) in Higher Education in Latin America

2.2 Generative Artificial Intelligence (GAI) in Higher Education in Latin America

Latin American countries are increasingly integrating Generative Artificial Intelligence (GAI) technology into their government strategies and economies. These initiatives will improve administrative efficiency, offering more personalized user services for the whole of society. However, the ill-intentioned use of (GAI) for process automation and decision-making presents many ethical and safety challenges, as it will compromise the security of society (Vinogradova, 2023). The application of GAI in higher education has shown potential to improve learning through the personalization of content and the automation of administrative processes (Guerrero-Quiñonez, et al. 2023).

However, the adoption of (GAI) is not without its challenges. Despite the digital transformation and the opportunities, it offers, concerns persist about data privacy, teaching automation, and the lack of specific regulation in the Latin American context (Andreoli, et al. 2024). While AI can improve learning and motivation, its growing use in Latin America presents challenges that demand proactive strategies to manage benefits, ethical concerns, and privacy issues. Therefore, the ability to expect medium- and long-term challenges is crucial to effectively address these changes (Rivera and Malaver, 2006). Reflecting on the past helps to interpret the present and manage desirable futures (Duque, Gonzáles and Santisteban, 2023).

2.3 Challenges of ChatGPT in Peruvian University Education

The language model developed by OpenAI, ChatGPT, is a Natural Language Processing (NLP) model capable of generating high-quality argumentative texts and engaging in realistic conversations. According to King (2023), it is the most powerful GAI platform that provides user-centric conversational experiences. ChatGPT has the potential to transform education by offering personalized learning experiences to improve students' language skills.

According to a recent study, 71.3% of Peruvian university students use ChatGPT as a support in their academic activities, improving their productivity and access to relevant information. Despite these benefits, its implementation in higher education requires a structured strategy to avoid excessive dependence and ensure its ethical use (Castillo, Palacios and Silva, 2023).

According to Velíbor and Indrasen (2023), ChatGPT should be used as a complement in learning and not as a substitute for critical thinking and human interaction in educational processes.

In addition, it is essential to establish clear guidelines on its application in teaching to avoid problems related to misinformation and the automatic generation of content without academic supervision (Rathore, 2023).

Many studies have addressed the application of Artificial intelligence in higher education at a global level (Zawacki-Richter et al., 2019). However, the literature on its impact in the Peruvian context remains limited. Most research focuses on countries with high technological infrastructure, leaving aside the specific challenges faced by universities in developing countries. In addition, the future scenarios and strategies necessary to ensure an ethical and effective implementation of AI in Peruvian universities have not been explored in depth. This study bridges that gap by applying a forward-looking approach to investigate how the integration of ChatGPT can transform Peruvian universities by 2033.

3. Methodology

The methodology is based on a theoretical foundation, which includes conceptual frameworks to initiate the stages of the research process.

3.1 Strategic Foresight

Foresight, initiated by Berger (1957), seeks to understand and influence the future through present decisions, creating possible futures. Foresight is a rigorous discipline that illuminates present actions with desirable futures. Integrating future perspectives into education is essential to help students guide and plan their personal and social projects, allowing them to imagine and work towards desirable and fair futures (Medina and Ortegón, 2006). Seeing the future as hopeful and full of possibilities allows young people to fully develop and be themselves (Duque, Gonzáles and Santisteban, 2023).

Strategic foresight analyzes the long-term future of science, economics, and society to identify emerging technologies that generate economic and social benefits (Gavigan, 2001). Iden, Methlie, and Christensen (2017) and the European Commission (2001) describe five rules that form the basis for foresight to achieve better results. Communication must be established with all the actors in the Research and Development (R+D) system involved in the prospective study. The focus should be on the long term to address forecast tasks. This must be followed by coordinated actions, consensus on the vision of the object of study and, ultimately, commitment to the results achieved.

3.2 Delphi Method

This tool is suitable for research because of its ability to structure the collection of expert responses. It has allowed the validation of impellers and has ensured the deliberation of experts in real time (Weaver, 1971). This method addresses complex problems through a structured group communication process based on rounds of surveys conducted by independent experts (Pätäri, 2010). This method facilitates collective contribution through

feedback, evaluation of group judgment, and review of individual opinion (Rowe and Wright, 1999). To ensure success, addressing challenges such as mitigating bias and encouraging expert participation is essential (Linstone and Murray, 1975). Although the future is unpredictable, some argue it can be estimated with some accuracy by people who have a good understanding of current conditions and can imagine possible future scenarios (Calleo & Pilla, 2023). The real-time Delphi method is an essential tool in technology forecasting. It is widely used in technological corporations and in the management of complex problems (Hirschberg and Rescher, 1960).

However, the Delphi method has weaknesses and limitations to consider. It is assumed that the experts participating in the Delphi have equivalent knowledge and experiences, which may not be true, as the knowledge of the panelists could be uneven. This could make the process more difficult (Chia-Chien and Sandford, 2019). The Delphi technique, performed in rounds, is time-consuming to complete the research process, as days and weeks can pass between rounds. While rounds improve the accuracy of results, they also increase the demand for more time to complete the process (Cunliffe, 2002). Therefore, for this study, the Delphi technique is applied in real time to optimize results, reduce biases and accelerate expert responses (Astigarraga, 2003), as well as to develop an active role of the researcher to ensure response rates and accuracy of results (Ludwig, 1997).

This article adopts a quantitative approach with a cross-sectional design, allowing for the objective collection of numerical data at a given time. This approach is ideal for measuring and analyzing experts' perceptions of ChatGPT's effects without manipulating variables, providing an accurate snapshot of the current context (Hernández, Fernández and Baptista, 2014; Hernández-Sampieri and Mendoza, 2018).

4. Results

The study aims to generate possible scenarios on the impact of ChatGPT on Peruvian universities by 2033. To achieve this objective, four stages were followed based on the foresight approach: (i) exploration of the system to identify drivers, (ii) validation of information to confirm drivers using the Delphi method in real time, (iii) construction of scenarios using Schwartz axes and structural analysis, and (iv) validation of scenarios through the Probability method, Desirability and Governance (PDG) and backcasting to develop strategies that lead to the target scenario towards 2033 (see Table 1).

A brief description of the current situation of Peruvian universities is presented to characterize the sample, noting that the number of students from various socioeconomic levels has increased in recent years. However, this process did not coincide with an improvement in the quality of the education offered (Ruíz-Gonzáles and Briceño-Cotrina, 2020). Peruvian universities face challenges, including improving the quality of teaching and applied research, and often lack a robust infrastructure. There are significant gaps that affect education, such as economic disparities that show a marked difference in infrastructure, budgetary resources, and higher quality education in the capital's universities compared to remote regions.

4.1 Data Analysis

The study selected a sample of 44 experts from public and private universities in Perú, the sample was chosen through non-probabilistic sampling (Hernández, Fernández and Baptista, 2014). Before distributing the surveys, an academic was consulted, and the pilot test was carried out with 10 experts. Based on their feedback, adjustments were made to improve the quality and accuracy of the instrument. To mitigate potential biases in expert input, multiple rounds of validation were conducted to ensure that the responses were transparent. In addition, three types of expert profiles were considered, which generated a variety of opinions and perspectives derived from homogeneous views, according to Winkler and Moser (2016). In the application of the Delphi survey in real time, three areas were established: education, technology and pedagogical management. The experts come from different professional fields: public sector (27.3%), private sector (50%), academia (13.6%), civil society (2.3%) and others (6.8%). The latter group includes experts with experience in related fields, whose perspective complements the interdisciplinary approach of this study. Regarding the experience of the experts, 47.7% have 2 to 5 years in the area of education, 36.4% have 6 to 10 years in technology, 9.1% have 11 to 15 years in pedagogical management and 6.8% have 16 or more years of experience in areas such as computer engineering, information economics and information technology. The samples size suffices to identify the relevant drivers and build prospective scenarios.

The research methodology follows the sequence outlined in the studies of Keenan and Popper (2008) and Medina and Ortegón (2006), as shown in Table 1.

Table 1: Stages of the prospective methodology process

Stages	Methodological tools
System Scan	Identifying drivers
Information Validation	Driver validation with the real-time Delphi survey app
Scenario construction	Schwartz axes and structural analysis
Scenario validation	Probability, Desirability and Governance (PDG) Method and Backcasting

4.2 System Scan

4.2.1 Identifying drivers

According to Ortega (2016), factors of change are driving variables that directly influence the subject of study, significantly shaping the construction of future scenarios. Twenty drivers are proposed for the prospective study of ChatGPT in Peruvian universities, based on bibliometric analysis and reviewed literature, as shown in Table 2. Each driver has a statement that serves as the basis for the construction of the Delphi survey in the later stage. These statements are derived from the influence of each driver. There are four areas that group the drivers: Educ (education area) which encompasses the first nine drivers, Tec (technology) which encompasses five drivers and GestP (pedagogical management) which encompasses six drivers.

Table 2: Drivers with their statements and references

Drivers	Declarations	References
Development of digital competences in university students	Students' digital skills will increase by 70%.	(Bozukarpan, Laoha and Jantakun, 2023)
Training students with critical thinking	The mastery of critical thinking will increase by 60%.	(García-Peñalvo, 2021)
Increased student academic performance	Students' academic performance will increase thanks to the use of digital technologies.	(Sánchez-Caballé, Gisbert-Cervera, and Esteve-Mon, 2020)
Feedback to students through intelligent tutoring	Cognitive feedback through intelligent tutoring will increase by 70%.	(Peters, et al., 2022)
Using ChatGPT for Academic Skills Development	Skills and abilities will increase with the use of ChatGPT by 70%.	(Fernández-Luque, Ramírez-Montoya, and Cordón-García, 2021)
Motivation	The use of ChatGPT will generate motivation and active participation in university teachers and students.	(Starkey, 2019)
Personalized and constructive education through ChatGPT	Personalized education through ChatGPT will increase by 70%.	(Colás-Bravo, Conde-Jiménez, and Reyes-de-Cózar, 2021)
Generation of scientific articles	The production of scientific articles involving teachers and students will increase by 40%.	(Harith et al., 2022)
Digital literacy of teachers in the teaching-learning process	The digital literacy rate of teachers and students will reach 70%, ensuring their competence in the effective use of digital technologies.	(Fernández-Batanero et al., 2021)
Use of ChatGPT in students with disabilities	The accessibility to the use of ChatGPT translates into a 40% increase in its usefulness for people with disabilities.	(Pérez-Jorge and Martínez-Murciano, 2022)
Integrating ChatGPT into Pedagogical Performance	The integration of ChatGPT in pedagogical performance will be 70%.	(Litiņa and Miltuze, 2023)
Technological infrastructure	There will be an advance in technological infrastructure by 2033, exceeding 50% implementation.	(Fernández-Batanero, et al., 2020)
Rethinking classroom methodology with ChatGPT	The teaching methodology in the classroom will improve with the incorporation of ChatGPT by 60%.	(Esteve-Mon, Llopis-Nebot, and Adell-Segura, 2020)

Drivers	Declarations	References
Use of ChatGPT to improve technological and training activities	There will be an improvement in technological and training activities, allowing them to perform more efficiently by 60%.	(Gutiérrez-Ángel et al., 2022)
Managing educational innovation using ChatGPT	Institutional educational innovation strategies will be designed at 50%.	(Ostanina et al., 2023)
Development of didactic skills in teachers	The development of teaching skills in teachers will be increased by 50%.	(Fernández-Luque, Ramírez-Montoya, and Cordón-García, 2021)
Use of ChatGPT for the design of class sessions and methodological strategies	The development of class sessions with the support of ChatGPT will be 60%.	(Martínez-Murciano, 2022)
Curricular integration according to new technological trends	Update curriculum design by promoting methodological innovations for the classroom and learning experiences.	(Litiņa and Miltuze, 2023)
Adapting careers to new technologies	80% of professional careers aligned with new technological trends will be implemented.	(Excluding Naumeca and Āboliņa, 2023)
Economic financing for the improvement of technological infrastructure	Funding will be obtained for the technological improvement of the university, plus 50%.	(Starkey, 2019)

4.3 Information Validation

4.3.1 Driver validation using the real-time Delphi survey

The real-time Delphi survey validates the drivers by consulting a group of experts individually through a questionnaire. This method aims to collect the opinion of experts on a topic related to future overtime (Rowe and Wright, 1999). The statements in the table above are part of the Delphi survey, composed of 20 statements derived from the influence of each driver. The survey was conducted using Google Forms, measuring importance based on three criteria: high, medium and low; and uncertainty based on five criteria as shown in Table 3. After conducting the Delphi survey with the 44 experts, the results are converted into percentage value to define our score. The results of the Delphi survey will facilitate the selection of the drivers that belong to quadrant III.

To determine the level of importance (high, medium, and low) and the level of uncertainty, the Pareto principle was applied, which states that 80% of problems stem from 20% of causes. This principle is used to improve reliability by identifying the components with the greatest need (Ortega, 2016). In this context, the high and medium levels were grouped, establishing a threshold of 80% or higher in both dimensions. This allowed the identification of the key drivers: critical thinking skills, cognitive feedback through intelligent tutoring and technological infrastructure, which were identified as the most important and uncertain and will serve as the basis for the construction of scenarios.

From the results of the survey, the level of importance (high, medium, low) of "critical thinking skills" indicates that 61.36% and 31.82% (grouped high and medium level together make more than 80%) of the experts consider it very important. These results exceed 80%, so it is considered "very important", that is, the level of response exceeds 80% due to its majority nature, which shows a significant consensus among the respondents. For the level of uncertainty, 43.18% of respondents believe that event could occur between 2023 and 2027, which does not exceed 80%, so it is considered "very uncertain". Similarly, the other two drivers (feedback to students through intelligent tutorials and technological infrastructure) with similar results are considered "very important and very uncertain". The level of importance has a total sum of 100%, referring to the total number of respondents (44 experts), and the same applies to the level of uncertainty.

Based on these criteria, key drivers with high importance and high uncertainty were selected. According to the results, these drivers are the most critical for the system under study and have a highly unpredictable behavior. Identifying and prioritizing these drivers is critical because they significantly influence the design of future scenarios. The three drivers that meet these characteristics are essential for strategic planning and decision-making, as their impact and volatility can determine several possible scenarios. These key drivers make it possible to anticipate and prepare responses to different situations, providing a solid foundation for the construction of robust and adaptive scenarios.

Table 3: Delphi Survey Results

Drivers of change	Importance (Level of importance of each statement)			*80 %*	Uncertainty Period during which the event is estimated to develop or occur)					*50%*
Mark each item (importance, experience, and uncertainty) with an X	High	Middle	Low	Importance	It's over	It will happen between 2023-2027	It will occur between 2028-2032	This will happen after 2033	That will never happen	Uncertain
Students' digital skills will increase by 70%.	63,64%	31,82%	4,55%	+	0,00%	54,55%	29,55%	11,36%	4,55%	-
The mastery of critical thinking will increase by 60%.	61,36%	29,55%	9,09%	+	0,00%	43,18%	40,91%	11,36%	4,55%	+
Students' academic performance will increase thanks to the use of digital technologies.	52,27%	40,91%	6,82%	+	6,82%	54,55%	20,45%	15,91%	2,27%	-
Cognitive feedback through intelligent tutoring will increase by 70%.	63,64%	27,27%	9,09%	+	2,27%	47,73%	36,36%	13,64%	0,00%	+
Skills and abilities will increase with the use of ChatGPT by 70%.	65,91%	34,09%	0,00%	+	9,09%	63,64%	22,73%	2,27%	2,27%	-
The use of ChatGPT will generate motivation and active participation in university teachers and students.	63,64%	27,27%	9,09%	+	2,27%	38,64%	52,27%	6,82%	0,00%	-
Personalized education through ChatGPT will increase by 70%.	34,09%	54,55%	11,36%	+	0,00%	54,55%	29,55%	9,09%	6,82%	-
The production of scientific articles involving teachers and students will increase by 40%.	63,64%	29,55%	6,82%	+	2,27%	50,00%	15,91%	31,82%	0,00%	-

Drivers of change	**Importance** (Level of importance of each statement)			***80 %***	**Uncertainty** Period during which the event is estimated to develop or occur)					***50%***
Mark each item (importance, experience, and uncertainty) with an X	**High**	**Middle**	**Low**	**Importance**	**It's over**	**It will happen between 2023-2027**	**It will occur between 2028-2032**	**This will happen after 2033**	**That will never happen**	**Uncertain**
The digital literacy rate of teachers and students will reach 70%, ensuring their competence in the effective use of digital technologies.	61,36%	22,73%	15,91%	+	0,00%	52,27%	38,64%	4,55%	4,55%	-
The accessibility to the use of ChatGPT translates into a 40% increase in its usefulness for people with disabilities.	68,18%	27,27%	4,55%	+	9,09%	56,82%	25,00%	9,09%	0,00%	-
The integration of ChatGPT in pedagogical performance will be 70%.	75,00%	20,45%	4,55%	+	2,27%	72,73%	18,18%	6,82%	0,00%	-
There will be an advance in technological infrastructure by 2033, exceeding 50% implementation.	59,09%	36,36%	4,55%	+	9,09%	43,18%	34,09%	11,36%	2,27%	+
The classroom teaching approach will be enhanced by incorporating ChatGPT.	77,27%	15,91%	6,82%	+	2,27%	50,00%	31,82%	15,91%	0,00%	-
There will be an improvement in technological and training activities, allowing them to perform more	77,27%	22,73%	0,00%	+	2,27%	59,09%	25,00%	11,36%	2,27%	-

Drivers of change	Importance (Level of importance of each statement)			*80 %*	Uncertainty Period during which the event is estimated to develop or occur)					*50%*
Mark each item (importance, experience, and uncertainty) with an X	**High**	**Middle**	**Low**	**Importance**	**It's over**	**It will happen between 2023-2027**	**It will occur between 2028-2032**	**This will happen after 2033**	**That will never happen**	**Uncertain**
efficiently by 60%.										
Institutional educational innovation strategies will be designed at 50%.	70,45%	25,00%	4,55%	+	2,27%	72,73%	20,45%	2,27%	2,27%	-
The development of teaching skills in teachers will be increased by 50%.	79,55%	18,18%	2,27%	+	2,27%	68,18%	20,45%	9,09%	0,00%	-
The development of class sessions with the support of ChatGPT will be 60%.	59,09%	36,36%	4,55%	+	0,00%	72,73%	22,73%	2,27%	2,27%	-
Update curriculum design by promoting methodological innovations for the classroom and learning experiences.	79,55%	18,18%	2,27%	+	4,55%	75,00%	15,91%	4,55%	0,00%	-
80% of professional careers aligned with new technological trends will be implemented.	75,00%	20,45%	4,55%	+	9,09%	52,27%	34,09%	2,27%	2,27%	-
Funding will be obtained for the technological improvement of the university, plus 50%.	61,36%	36,36%	2,27%	+	6,82%	54,55%	29,55%	9,09%	0,00%	-

4.4 Scenario Construction

4.4.1 Schwartz Axes

The level of importance and uncertainty are fundamental in the 2x2 matrix to select drivers and build scenarios. Importance refers to how essential a driver is to the studied system under study, while uncertainty relates to the degree of unpredictability associated with its behavior (van't Klooster and van Asselt, 2006). The drivers in quadrant III (+ important and + uncertain) significantly influence the development of various scenarios, making them key drivers with the most significant future impact.

Axes are used to classify factors based on their level of importance and uncertainty. Quadrant I, labeled "Environment", contains the least uncertain and most important factors. The most important and least uncertain factors are placed in quadrant II, "Baseline". The more uncertain and more important factors are identified in quadrant III, called Diversity. The elements in this quadrant create diversity or differences between scenarios and are, therefore, the most commonly used for scenario construction. The factors in quadrant III form the most important and probable basis for the construction of scenarios. Finally, the least uncertain and most important factors are located in quadrant IV, "Detail", and only affect a few scenarios (Schwartz, 1991).

Table 4 shows that the drivers "Training students with critical thinking skills, Feedback to students through intelligent tutorials and Technological infrastructure" present high levels of importance and uncertainty, belonging to quadrant III.

Table 4: Ranking Key Factors by Quadrant

Identification	Conductor	Importance	Uncertainty	Quadrant
Educ2	Training students with critical thinking skills	+	+	III
Educa4	Feedback to students through intelligent tutoring	+	+	III
Tec3	Technological infrastructure	+	+	III

These drivers are grouped into two axes: the first axis is called "Promotion of critical thinking", which includes two drivers, and the second axis is called "State-of-the-art technological infrastructure", as shown in Table 5.

Table 5: Drivers of change and axes of uncertainty.

Identification	Drivers of change	Axes of uncertainty
Educ2	Training students with critical thinking skills.	Fostering critical thinking
Educ4	Feedback to students through intelligent tutorials.	
Tec3	Technological infrastructure.	State-of-the-art technological infrastructure

Identifying the driving factors is crucial to building favorable scenarios, such as:

- Training students with critical thinking skills is essential for academic and professional development. This factor will drive the need for innovative pedagogical methods that stimulate reflection and analysis in students, preparing them to face the cognitive challenges to come (Glen, 1995).
- Personalized feedback from intelligent mentoring is essential to improving the quality and accessibility of education. It will allow students to adapt to the learning process, meeting individual needs and fostering various skills. It also improves the effectiveness of the educational process by providing personalized support to each student, resulting in more comprehensive and meaningful learning (Kulik and Fletcher, 2016).
- Technology infrastructure is critical to preparing students for a digital world. It will provide tools and resources to improve teaching and learning, enabling access to diverse information and facilitating collaboration. Key factors include the implementation of digital classrooms, training teachers in the effective use of technology, and investment in digital educational resources (Lamb and Weiner, 2021).

4.4.2 Structural analysis

The structural analysis examines the relationship of dependence and influence between the elements of quadrant III of the Schwartz axis, providing a visual representation of their impact on the construction of future scenarios, as shown in Figure 1.

According to Table 6 the correlation coefficient r = + 0.85 indicates that there is a positive relationship since the greater the dependence, the greater the influence of the drivers Furthermore, 71.71% of the variability of "influence" is explained by "dependence" in the regression model, suggesting a moderate to strong relationship between both variables.

In this sense, the"Educ2" (Educating students with critical thinking) engine of change is highly dependent on other factors, which indicates that by itself this variable does not drive large actions, which leads to having a low impact. "Educ4" (Intelligent Tutoring) has a medium level of dependency, indicating that feedback through tutoring is influenced by various factors, such as barriers in the Peruvian context or limitations in teacher training. However, it does not cease to have a significant impact on the system. Finally, "Tec3" (Technological Infrastructure) has a low dependency, so it does not depend on other factors, therefore it is a key factor that serves as an effective strategy to strengthen the educational system.

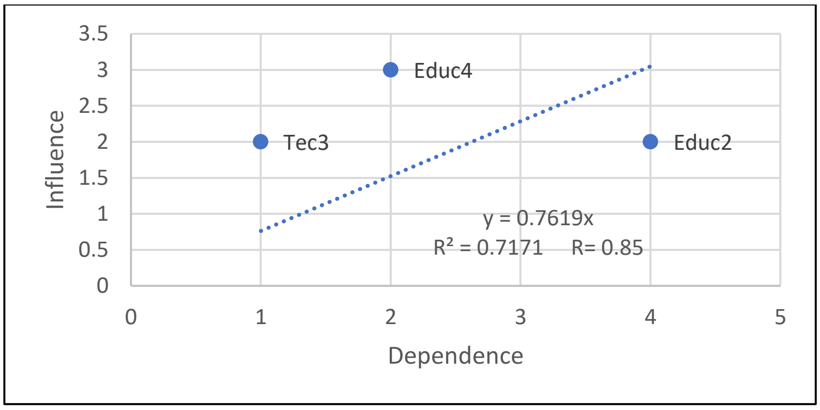

Figure 1: Relationship of structural analysis in the plane of dependence-influence of the drivers

Table 6: Structural Analysis Matrix

	Educ2	Educ4	Tec3	∑ Dependency
Educ2	---	2	2	4
Educ4	2	----	0	2
Tec3	0	1	----	1
∑ Influence	2	3	2	5

In Table 7, the scenarios are derived from the combination of positive and negative signs (+ -) that reflects a behavior that enhances or decreases the elements represented on each axis. Positive values (+) indicate behavior that emphasizes the characteristics of the elements, while negative values (-) indicate actions that reduce their impact. With two axes of uncertainty, four different scenarios are generated. Once the possible scenarios have been described, an internal consistency analysis is performed to ensure that there are no conflicts between the combinations of poles in each scenario. According to the analysis carried out, it is observed that all scenarios have consistency with the combination of critical factors.

Table 7: Scenarios derived from structural analysis

Scenarios	Axis III: "Promoting critical thinking"	Axis III "State-of-the-art technological infrastructure"	Stage Name	Consistency Analysis
1	+	+	Digital mentoring and integrated technology	Coherent
2	+	-	Technological advances and challenges	Coherent
3	-	+	Overcoming challenges with innovation and technology	Coherent
4	-	-	Educational Challenges 2033	Coherent

4.4.3 Scenario design

For the generation of the scenarios, the key drivers of the result of the Delphi Method are considered, as they have high levels of importance and uncertainty, belonging to quadrant III of the 2x2 matrix, including 'Training of students with critical thinking skills, Feedback to students through intelligent tutoring and Technological infrastructure'. Based on these 3 drivers, the scenarios are generated.

Scenario 1: Digital mentoring and integrated technology

This scenario reflects a genuine commitment to developing individuals with critical thinking skills essential to success in an ever-changing world. Intelligent tutoring adapted to the specific needs of each student represents a significant advance in personalized education, allowing for more effective and comprehensive academic development. Tightly integrated into the educational environment, advanced technological infrastructure democratizes access to high-quality digital resources, reducing learning gaps and promoting educational equity. This combination of factors represents a realistic evolution towards a more inclusive and dynamic education system, preparing students to face the challenges of the 21st century with confidence and competence.

Scenario 2: Technological advances and challenges

In the projected scenario towards 2033, the educational approach prioritizes the development of critical thinking and the implementing intelligent tutorials to provide personalized and effective feedback to students. However, important challenges related to technological infrastructure limit the full integration of digital tools into the educational process. These obstacles can generate gaps in access to digital education and affect the quality of learning. Despite these challenges, it is important to note that the educational landscape shows crucial advances for a more complete and fair development. There are opportunities to address technological challenges and improve the effectiveness of the education system, promoting a dynamic and adaptive learning environment for all students.

Scenario 3: Overcoming challenges with innovation and technology

The year 2033 presents important challenges in the training of students in critical thinking and personalized tutoring. However, technological infrastructure has become a key strength, providing advanced tools that improve connectivity and access to digital resources. Integrating technology in higher education has also provided opportunities for international collaboration and knowledge sharing, democratizing access to education and allowing people from diverse backgrounds to take part in educational programs, all under the development of an ethical regulatory framework.

Scenario 4: Educational challenges towards 2033

In this scenario, the development of critical thinking is hindered by a traditional educational approach that prioritizes memorization over analysis and reflection. Feedback to students needs to improve because of a lack of resources and adequate training for educators, which limits the quality and personalization of feedback. Technological infrastructure experiences significant setbacks because of a lack of investment and upgrades, resulting in limited access to digital tools and an outdated learning environment. These factors combined create a challenging educational landscape that is less conducive to the integral development of students.

4.5 Scenario Validation

4.5.1 Probability, Desirability, and Governance (PDG) method

In this section, scenario validation was carried out with the participation of 16 experts, using the criteria of probability, desirability, and governance for each scenario. Each scenario was assigned values according to the aforementioned criteria using a Likert scale, where 1 represents less likely, desirable or governable, and 4 represents more likely, desirable or governable. The highest total value determines the target scenario for strategy development. According to the results, Scenario 3, "Overcoming challenges with innovation and technology", was identified as the target scenario, as shown in Table 8.

Table 8: Results of the criteria of probability (P), desirability (D) and governance (G)

Scenarios	P	D	G	Total
1: Digital mentoring and integrated technology	36	72	44	152
2: Technological advances and challenges	53	52	54	158

Scenarios	P	D	G	Total
3: Overcoming challenges with innovation and technology	63	54	59	176
4: Educational challenges towards 2033	45	26	48	119

4.5.2 Retrospection: Strategy formulation

The retrospection method was used to develop strategies leading to the target scenario (Höjer and Mattsson, 2000). Retrospection is a methodology that starts with a desired future goal and works backwards to identify the actions needed to achieve it. It provides a solid framework for strategic planning and long-term policy development (Soria-Lara and Banister, 2017). This approach helps to visualize a desired future and develop retroactive strategies to reach the goal, involving multiple stakeholders in the planning process. Table 9 outlines strategies aimed at addressing digital challenges and promoting the effective adoption of technologies such as ChatGPT in universities, maximizing their benefits to enhance educational quality.

Table 9: Strategy formulation: retrospection

Year	Milestone	Strategy
2023	Present	Organize interactive workshops and seminars for educators and students aimed at enhancing the benefits of integrating ChatGPT into education. Establish student and faculty engagement committees to gather opinions and concerns, fostering a collaborative approach to decision-making.
2024	Milestone 5: Needs and resource assessment	Conduct a comprehensive assessment of the institution's technological needs. Hire consultants specializing in educational technology to evaluate existing infrastructure and propose improvements.
2026	Milestone 4: Strategic investment	Acquire and implement advanced educational software and tools that facilitate personalized feedback and the development of critical thinking. Establish ongoing training programs for teachers and students, focusing on the effective use of ChatGPT and addressing initial concerns about the loss of analytical skills. Organize training courses in educational technology and digital ethics that encourage responsible and effective use of this algorithm.
2028	Milestone 3: Continuous monitoring and adjustment	Establish a monitoring system to assess the impact of integrating ChatGPT into education. Develop a curricular plan where the use of emerging technologies is incorporated into the teaching-learning process, with a proactive approach by adapting a practical methodology in teaching. Continuously adjust infrastructure and training programs based on feedback and results achieved.
2030	Milestone 2: Developing innovative educational content	Collaborate with pedagogical and technological experts to develop innovative educational content that maximizes ChatGPT's capabilities. Integrate this content into the curriculum to enhance educational experience.
2032	Milestone 1: Expansion of training programs	Expand training programs to include not only teachers and students but also parents and members of the educational community. Evaluate the success of your ChatGPT implementation, analyzing metrics such as academic performance, student engagement, and teacher feedback. Continue to foster a culture of continuous learning and adaptability, preparing the educational community to face future technological changes.
2033	Meta Scenario: Overcoming challenges with innovation and technology	Achieve the target scenario.

The retrospection method was used to develop strategies leading to the target scenario (Höjer and Mattsson, 2000). Retrospection is a methodology that starts with a desired future goal and works backwards to identify the actions needed to achieve it. It provides a solid framework for strategic planning and long-term policy development (Soria-Lara and Banister, 2017). This approach helps to visualize a desired future and develop

retroactive strategies to reach the goal, involving multiple stakeholders in the planning process. Table 9 outlines strategies aimed at addressing digital challenges and promoting the effective adoption of technologies such as ChatGPT in universities, maximizing their benefits to enhance educational quality.

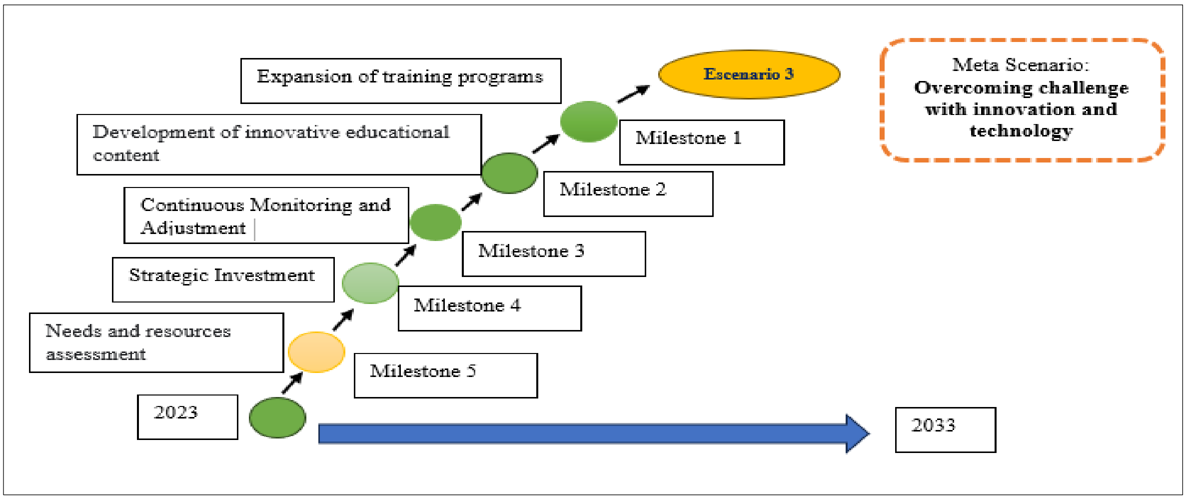

Figure 2: Temporal milestones in backcasting

5. Discussion and Conclusions

5.1 Discussion

The findings of this study highlight the importance of adapting to the ever-evolving academic environment by incorporating digital tools, such as ChatGPT, fused with human intellect. This integration will lead to a profound change in the dynamics of academic performance.

According to Castillo, Palacios and Silva (2023), although the use of ChatGPT improves students' productivity and access to relevant information, its implementation in higher education requires clear strategies to avoid over-reliance and ensure its ethical use. Velíbor and Indrasen (2023) argue that this tool should be used as a complement to learn, without replacing critical thinking and human interaction, an aspect that reinforces the need for an ethical framework that regulates its use.

In addition, Guadalupe et al. (2023) highlight that, while ChatGPT offers opportunities to personalize learning, the risk of misinformation and bias persists, requiring responsible and regulated integration. This research contributes to the debate by proposing ethical management strategies, aligning with Andreoli et al. (2024), who underscore the importance of addressing privacy and data security concerns to protect students and educators.

This research shows that integrating ChatGPT in Peruvian universities represents a significant opportunity to innovate in the classroom. This is affirmed in the outcome of the target scenario, where the integration of critical thinking skills and feedback through intelligent personalized tutorials is strengthened by an advanced technological infrastructure. Democratizing access to education and allowing the participation of students and teachers from various backgrounds in academic programs, all under an ethical framework where a responsible use of artificial intelligence in the academic field is guaranteed.

In the context of Peru's digital transformation, Law N°. 31814 was introduced to promote the use of artificial intelligence through a regulatory framework that ensures its responsible, ethical, and transparent implementation in both public and private sectors (Gobierno de Perú, 2025). However, its implementation faces challenges related to infrastructure and personnel training, which are essential for the proper use of these technologies.

The study advocates for the implementation of management strategies and ethical guidelines that uphold academic integrity and prioritize the well-being of students and educators, ensuring that tools like ChatGPT are used appropriately and in alignment with existing regulations.

- Ethical risks. - Risks in the educational environment are related to academic integrity and indiscriminate access to these tools (Peng and Zhao, 2024) foster academic dishonesty, by allowing students to generate information without any control (Cotton, Cotton and Shipway, 2023). The research recommends that the university at the academic level adopt a proactive and ethical approach in the use of this tool, that is, that they make changes in the curriculum, incorporating courses on digital ethics and promoting the responsible use of artificial intelligence.

- Ethical barriers. - Algorithmic biases are a barrier in using ChatGPT, since it can influence the generation of responses presenting inequalities and affecting the clarity and objectivity of the information (Ferrara, 2023). In the literature review of this research, Velibor and Indrasen (2023) argues that this problem can be minimized, as long as ChatGPT is used as a complement and not as a substitute for critical thinking and human interaction. Therefore, it is essential to adopt a balanced approach that integrates the use of AI with critical thinking analysis, ensuring a more objective and ethical use of technology.
- Concerns. - In contrast to the review of the literature, several authors agree that the ethical and educational concerns derived from the implementation of the Constitution. According to Guadalupe et al. (2023), although ChatGPT provides many opportunities to personalize learning, there is a risk of reproducing biases and misinformation, which affects the quality of information.

On the other hand, according to the theory, Kuleto et al. (2021) shows where they warn about the impact on the teacher's role, pointing out that automation can reduce pedagogical interaction and affect personalized teaching.

5.2 Conclusion

The identification and validation of the 20 drivers through the Delphi method in real time have confirmed the importance and relevance of influencing ChatGPT in Peruvian universities. This process has provided a solid foundation for decision-making and the development of scenarios that effectively integrate ChatGPT into the educational process.

The validation of these drivers supports the understanding of the key elements that influence the successful implementation of ChatGPT in Peruvian universities.

The selection of the driving factors, such as the training of students in critical thinking, feedback through intelligent tutorials and the development of technological infrastructure, is closely related to the influence of ChatGPT on Peruvian universities by 2033.

However, it is important to mention that implementing milestone retrospection strategies must be a structured process that requires adaptive and critical management. In this regard, funding for its implementation is crucial, as without sufficient technological and financial resources, the strategies may be inadequate. Resistance to digital change represents an obstacle that must be managed with a comprehensive approach. In addition, it is important to recognize and overcome digital literacy gaps, which can exacerbate inequalities in other universities that do not have the same resources.

The target scenario highlights the advances and challenges that could arise in the way of more effective and technologically advanced education. In addition, it stresses the importance of addressing ethical aspects in using tools such as ChatGPT, promoting its responsible implementation to guarantee academic integrity, protect user privacy, and foster authentic learning, where technology acts as support and not as a substitute for critical thinking and personal effort.

Technology, applied with a solid plan and commitment to quality education, can be key to learning and personal development. The goal for 2033 is to overcome challenges through innovation and technology, with awareness-raising, feedback and continuous training programs. It is essential to implement effective technological strategies and to constantly monitor educational content and the development of innovative programs.

Future research may explore other generative AI algorithms, to provide a deeper and more comprehensive understanding of the use of these tools in academia.

Statement Of Artificial Intelligence: An artificial intelligence tool was used specifically to support grammar correction and proofreading in the translation of the document. The content generated with the tool was edited and validated by the authors.

Statement Of Ethics: This article does not require ethical approval, as no human subjects were involved.

Future Lines of Research: This research lays the groundwork for numerous promising lines of future study. In view of the dynamic nature of AI technologies and their potential influence on higher education, two key areas deserve further exploration:

Longitudinal studies are essential for tracking the evolutionary impact of ChatGPT and similar GAI technologies on universities, as they provide empirical data that can refine forecasting scenarios and strategies. These studies allow researchers and policymakers to compare actual results with predictions. In addition, future research should address the ethical implications of AI in education, focusing on privacy, data security, and its role in widening or narrowing the digital divide. Addressing these issues is crucial to developing ethical guidelines and policies that protect both students and educators.

References

Ali, S., Ravi, P., Williams, R., DiPaola, D., & Breazeal, C. (2024). Constructing Dreams Using Generative AI. *Proceedings of the AAAI Conference on Artificial Intelligence, 38*(21), 23268–23275. https://doi.org/10.1609/aaai.v38i21.30374

Andreoli, S., Perillo, L., Aubert, E., & Cherbavaz, M. C. (2024). Entre humanos y algoritmos: percepciones docentes sobre la exploración con IAG en la Enseñanza del Nivel Superior. *Revista Iberoamericana de Tecnología En Educación y Educación En Tecnología, 37*, e6. https://doi.org/10.24215/18509959.37.e6

Andreucci-Annunziata, P., Riedemann, A., Cortés, S., Mellado, A., del Río, M. T., & Vega-Muñoz, A. (2023). Conceptualizations and instructional strategies on critical thinking in higher education: A systematic review of systematic reviews. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1141686

Anohina-Naumeca, A., & Āboliņa, A. (2023). Conceptualising Digital Competency for ICT Specialists. *Human, Technologies and Quality of Education*. https://doi.org/10.22364/htqe.2023.34

Astigarraga, E. (2003). El Método Delphi. Universidad de Deusto, pp. 1-14. [en línea] Disponible en: < http://prospectiva.eu/zaharra/Metodo_delphi.pdf > [Consultado el 15 de julio del 2024]

Berger, G. (1957). Sciences humaines et prevision. *Revuedes Deux Mondes, (3)*. Recuperado de http://www.revuedesdeuxmondes.fr/archive/article.php?code=14758

Bojukrapan, S., Laoha, R., & Jantakoon, T. (2023). Synthesizing Digital Teacher Competencies for Teaching Profession Students in Higher Education. *Higher Education Studies*. https://doi.org/10.5539/hes.v13n4p227

Borjas, C., 2005. LEY DE PARETO APLICADA A LA FIABILIDAD., 8, págs. 1-9.Brunner, J., & Tadesco, J. (2003). Las nuevas tecnologías y el futuro de la educación. Ideas, Personas y Políticas/ UNESCO. [en línea] Disponible en: <https://educacion.udd.cl/files/2017/03/MS_Brunner-Educacion_escenarios_de_futuro-Nuevas_tecnologias_y_la_soc_info-1.pdf > [Consultado el 17 de diciembre del 2023]

Buchholz, K. (2023). ChatGPT sprints to one million users. Https://Www.Statista.Com/Chart/29174/Time-to-One-Million-Users/

Calleo, Y., & Pilla, F. (2023). Delphi-based future scenarios: A bibliometric analysis of climate change case studies. *Futures, 149*, 103143. https://doi.org/10.1016/j.futures.2023.103143

Castillo, C., Palacios, C. and Silva, M. (2023) *Tecnología educativa en América Latina y el Caribe*. SUMMA. [en línea] Disponible en : < https://www.summaedu.org/wp-content/uploads/2024/02/Tecnologia-educativa-LAC-Castillo-et-al-2023.pdf > [Acceso el 09 de setiembre del 2024]

Chia-Chien, H., & Sandford, B. (2019). The Delphi Technique: Understanding Consensus. *Practical Assessment, Research, and Evaluation, 12*. [in line] Available in: <https://openpublishing.library.umass.edu/pare/article/id/1418/> [Accessed 11 January 2024].

Colás-Bravo, P., Conde-Jiménez, J., & Reyes-de-Cózar, S. (2021). Sustainability and digital teaching competence in higher education. *Sustainability, 13*(22), 12354. https://doi.org/10.3390/su132212354

Cotton, D. R. E., Cotton, P. A. and Shipway, J. R. (2023) 'Chatting and cheating: Ensuring academic integrity in the era of ChatGPT', *Innovations in Education and Teaching International*, 61(2), pp. 228–239. doi: 10.1080/14703297.2023.2190148.

Cunliffe, S. (2002). Forecasting risks in the tourism industry using the Delphi. *Tourism, 50*(1), 31–41. [in line] Available in: < http://worldwideuniversity.org/library/The_Delphi_Technique_Making_Sense_of_Consensus.pdf> [Accessed 15 november 2024].

Duque, L. F., Gonzáles, G. A., & Santisteban, A. (2023). What Do You Think About the Future? Students' Imaginaries in Colombian Post-Conflict. *Journal of Futures Studies, 27*(4), 13–27. https://doi.org/10.6531/JFS.202306_27(4).0002

Esteve-Mon, F. M., Llopis-Nebot, M. Á., & Adell-Segura, J. (2020). Digital teaching competence of university teachers: A systematic review of literature. *Revista de Estudios de Futuros, 4*(1), 123-145. https://doi.org/10.3390/educsci1010004

European Commission. (2001). *A practical guide to regional foresight*. Foresight. https://www.foresight.pl/assets/downloads/publications/eur20128en-APracticalGuidetoRegionalForesight2001.pdf

Fernández-Batanero, J. M., Montenegro-Rueda, M., Fernández-Cerero, J., & García-Martínez, I. (2020). Digital competences for teacher development: A systematic review. *Revista de Estudios de Futuros, 3*(2), 76-89. https://doi.org/10.3390/educsci1010004

Fernández-Batanero, J., Román-Graván, P., Montenegro-Rueda, M., & López-Meneses, E. (2021). Digital teaching competence of university teachers: A systematic review of literature. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 15*, 399-406. https://doi.org/10.1109/RITA.2020.3033225

Fernández-Luque, A. M., Ramírez-Montoya, M., & Cordón-García, J. (2021). Training in digital competencies for health professionals: Systematic mapping (2015-2019). *Profesional de la información*. https://doi.org/10.3145/EPI.2021.MAR.13

Ferrara, E., 2023. ¿ChatGPT debería estar sesgado? Desafíos y riesgos de sesgo en modelos de lenguaje grandes. *First Monday* , 28. https://doi.org/10.5210/fm.v28i11.13346 .

García-Peñalvo, F. J. (2021). Transformación digital en las universidades: Implicaciones de la pandemia de la COVID-19. *Education in the Knowledge Society (EKS), 22*, e25465. https://doi.org/10.14201/eks.25465

García-Peñalvo, F. J. (2023). La percepción de la Inteligencia Artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge Society (EKS), 24*, e31279. https://doi.org/10.14201/eks.31279Graham, M. (2009). Scenarios: Is it worth the effort?. Journal of Futures Studies, 13(3), 81-92. [in line] Available in: https://jfsdigital.org/articles-and-essays/2009-2/vol-13-no-3-february/ [Accessed 1 september 2024].

Gavigan, J. (2001). Foresight for Regional Development Network. A Practical Guide to Regional Foresight. [in line] Available in: <https://www.foresight.pl/assets/downloads/publications/eur20128en-APracticalGuidetoRegionalForesight2001.pdf > [Accessed 5 august 2024].

Guerrero-Quiñonez, A.J., Bedoya-Flores, M.C., Mosquera-Quiñonez, E.F. and Mesías-Simisterra, Á.E. (2023) 'Artificial Intelligence and its scope in Latin American higher education', *Ibero-American Journal of Education & Society Research*, 3(1), pp. 264-271. doi: 10.56183/iberoeds. v3i1.627.

Glen, S. (1995). Developing critical thinking in higher education. *Nurse Education Today, 15*(3), 170–176. https://doi.org/10.1016/S0260-6917(95)80102-2.

Gobierno de Perú. (2025). *Ley N° 31814*. Congreso de la República del Perú. [en línea] Disponible en: <https://www.gob.pe/institucion/congreso-de-la-republica/normas-legales/4565760-31814> [Acceso 17 febrero 2025].

Guadalupe, A., Castillo, R., Serna, G., Arocutipa, J., Berrios, H., Antonio, M., Rodríguez, M., Yanowsky, G., López, H., Marina, R., Teves, V., Victor, H., Rivera, H. and Arias-Gonzáles, J., 2023. *Effect of ChatGPT on the digitized learning process of university students*. Journal of Namibia Studies: History, Politics, Culture. [online] Available at: https://doi.org/10.59670/jns.v33i.411 [Accessed 22 May 2025].

Gutiérrez-Ángel, N., Sánchez-García, J. N., Mercader-Rubio, I., García-Martín, J., & Brito-Costa, S. (2022). Digital literacy in the university setting: A literature review of empirical studies between 2010 and 2021. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.896800

Harith, S., Backhaus, I., Mohbin, N., Ngo, H., & Khoo, S. (2022). Effectiveness of digital mental health interventions for university students: An umbrella review. *PeerJ, 10*. https://doi.org/10.7717/peerj.13111

Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la investigación*. México: McGraw-Hill.

Hernández-Sampieri, R., & Mendoza, C. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. Ciudad de México, México: McGraw-Hill Education. [en línea] Disponible en: <https://apiperiodico.jalisco.gob.mx/api/sites/periodicooficial.jalisco.gob.mx/files/metodologia_de_la_investigacion_-_roberto_hernandez_sampieri.pdf > [Acceso 6 julio 2024].

Hirschberg, H., & Rescher, O. N. H. (1960). On the Epistemology of the Inexact Sciences. RAND Corporation. https://www.rand.org/pubs/reports/R353.html

Höjer, M., & Mattsson, L.-G. (2000). Determinism and backcasting in future studies. *Futures, 32*(7), 613–634. https://doi.org/10.1016/S0016-3287(00)00012-4

Iden, J., Methlie, L. B., & Christensen, G. E. (2017). The nature of strategic foresight research: A systematic literature review. *Technological Forecasting and Social Change, 116*, 87-97. https://doi.org/10.1016/J.TECHFORE.2016.11.002

Inayatullah, S., 2008. Seis pilares: pensamiento de futuro para la transformación. *Foresight*, 10, págs. 4-21. https://doi.org/10.1108/14636680810855991 .Keenan, M., & Popper, R. (2008). Comparing foresight "style" in six world regions. *Foresight, 10*(6), 16–38. https://doi.org/10.1108/14636680810918568

King, M. R. (2023). A Conversation on Artificial Intelligence, Chatbots, and Plagiarism in Higher Education. *Cellular and Molecular Bioengineering, 16*(1), 1–2. https://doi.org/10.1007/s12195-022-00754-8

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O., Păun, D., & Mihoreanu, L., 2021. Exploring Opportunities and Challenges of Artificial Intelligence and Machine Learning in Higher Education Institutions. *Sustainability*. https://doi.org/10.3390/su131810424.Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems. *Review of Educational Research, 86*(1), 42–78. https://doi.org/10.3102/0034654315581420

Lamb, A. J., & Weiner, J. M. (2021). Technology as infrastructure for change: district leader understandings of 1:1 educational technology initiatives and educational change. *Journal of Educational Administration, 59*(3), 335–351. https://doi.org/10.1108/JEA-10-2020-0220

Linstone, H. A., & Murray, T. (1975). The delphi method: techniques and applications. Addison-Wesley Pub. Co. Advanced Book Program. [in line] Available in: https://www.researchgate.net/publication/237035943_The_Delphi_Method_Techniques_and_Applications [Accessed 23 december 2023].

Ludwig, B. (1997). Predicting the Future: Have you considered using Delphi Methodology? Journal of Extension, 35. [in line] Available in: <https://archives.joe.org/joe/1997october/tt2.php > [Accessed 7 september 2024].

Luna, G., 2023. Estudio sobre el impacto de herramientas de inteligencia artificial en el desarrollo de clases universitarias en la escuela de comunicación de la Universidad Nacional José Faustino Sánchez Carrión. *Metaverse Basic and Applied Research* . https://doi.org/10.56294/mr202351 .Martínez-Murciano, M. C. (2022). Gamification with Scratch or App Inventor in higher education: A systematic review. *Future Internet, 14*(12), 374. [in line] Available in: https://doi.org/10.3390/fi14120374

Medina, J., & Ortegón, E. (2006). Manual de prospectiva y decisión estratégica: bases teóricas e instrumentos para América Latina y el Caribe. Instituto Latinoamericano y Del Caribe de Planificación Económica y Social (ILPES). [en línea] Disponible en: < https://repositorio.cepal.org/bitstream/handle/11362/5490/1/S0600190_es.pdf> [Consultado el 15 de noviembre del 2023]

Ortega, F. (2016). PROSPECTIVA EMPRESARIAL. Manual de corporate foresight para América Latina (1era edición). Universidad de Lima Fondo Editorial. [en línea] Disponible en: < https://www.academia.edu/116089308/Prospectiva_Empresarial_Manual_de_Corporate_Foresight> [Consultado el 8 de julio del 2023]

Ostanina, A., Bazyl, O., Tsviakh, O., & Dovzhuk, N. (2023). Formation of digital competence in higher education students as a basis for the transformation of education of the future. *Futurity Education*. https://doi.org/10.57125/fed.2023.25.03.10

Pätäri, S. (2010). Industry- and company-level factors influencing the development of the forest energy business — insights from a Delphi Study. *Technological Forecasting and Social Change, 77*(1), 94–109. https://doi.org/10.1016/j.techfore.2009.06.004

Peng, L., y Zhao, B., 2024. Navegando por el panorama ético detrás de ChatGPT. *Big Data Soc.* , 11. https://doi.org/10.1177/20539517241237488 .

Pérez-Jorge, D., & Martínez-Murciano, M. C. (2022). Gamification with Scratch or App Inventor in higher education: A systematic review. *Future Internet, 14*(12), 374. https://doi.org/10.3390/fi14120374

Peters, M., Elasri Ejjaberi, A., Martínez, M. J., & Fàbregues, S. (2022). Teacher digital competence development in higher education: Overview of systematic reviews. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.7543

Qawqzeh, Y., 2024. Exploring the Influence of Student Interaction with ChatGPT on Critical Thinking, Problem Solving, and Creativity. *International Journal of Information and Education Technology*. https://doi.org/10.18178/ijiet.2024.14.4.2082.

Rathore, B. (2023). Future of AI and Generation Alpha: ChatGPT beyond Boundaries. International Multidisciplinary Refereed Journal, 12. [in line] Available in: < https://www.eduzonejournal.com/index.php/eiprmj/article/view/254 > [Accessed 20 july 2023]

Rivera, H., & Malaver, M. (2006). La importancia de la prospectiva en la sociedad. *Universidad & Empresa, 5*(10), 257–270. [en línea] Disponible en: https://www.redalyc.org/pdf/1872/187217412011.pdf [Consultado el 15 de setiembre del 2024]

Rivas, A., Buchbinder, N. and Barrenechea, I. (2023) *El futuro de la inteligencia artificial en educación en América Latina*. España: ProFuturo y OEI, p. 48. Available at: https://www.ipn.mx/assets/files/innovacion/docs/Innovacion-Educativa/Innovacion-Educativa-96/el-futuro-de-la-inteligencia-artificial-ex-libris.pdf (Accessed: [Consultado el 5 de diciembre del 2023]Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting, 15*(4), 353–375. https://doi.org/10.1016/S0169-2070(99)00018-7

Ruíz-Gonzáles, C., & Briceño-Cotrina, O. (2020). Realidad y perspectiva de la Educación Superior en el Perú. *Revista Ciencia y Tecnología, 16*(4), 97–108. https://doi.org/10.17268/rev.cyt.2020.04.09

Sánchez-Caballé, A., Gisbert-Cervera, M., & Esteve-Mon, F. M. (2020). The digital competence of university students: A systematic literature review. *ALOMA, 38*(1), 63-74. https://doi.org/10.51698/ALOMA.2020.38.1.63-74

Sallam, M. (2023). ChatGPT Utility in Healthcare Education, Research, and Practice: Systematic Review on the Promising Perspectives and Valid Concerns. Healthcare, 11(6), 887. https://doi.org/10.3390/healthcare11060887

Schwartz, P. (1991). The Art of the Long View: Planning for the Future in an Uncertain World. Doubleday. [in line] Available in: < https://www.wiley.com/en-gb/Art+of+the+Long+View:+Planning+for+the+Future+in+an+Uncertain+World-p-9780471977858> [Accessed 1 december 2023].

SITEAL, & UNESCO. (n.d.). Perfil de país-Perú. [en línea] Disponible en: <Https://Siteal.Iiep.Unesco.Org/País/Perú> [Consultado el 9 octubre 2023].

Snijders, D., van der Duin, P., Marchau, V., & van Doorn, G. J. (2018). Scenarios for ICT-related education: A qualitative meta-analysis. *Journal of Futures Studies, 23*(2), 13-28. https://doi.org/10.6531/JFS.201812_23(2).0002

Soria-Lara, J. A., & Banister, D. (2017). Participatory visioning in transport backcasting studies: Methodological lessons from Andalusia (Spain). *Journal of Transport Geography, 58*, 113–126. https://doi.org/10.1016/j.jtrangeo.2016.11.012

Starkey, L. (2019). A review of research exploring teacher preparation for the digital age. *Cambridge Journ*Anohina-Naumeca*al of Education, 50*, 37-56. https://doi.org/10.1080/0305764X.2019.1625867

Van 't Klooster, S. A., & van Asselt, M. B. A. (2006). Practising the scenario-axes technique. *Futures, 38*(1), 15–30. https://doi.org/10.1016/j.futures.2005.04.019

Velíbor, B., & Indrasen, P. (2023). Chat GPT and Education. Preprint. [in line] Available in: https://www.researchgate.net/publication/369926506_Chat_GPT_and_education [Accessed 10 july 2023].

Vidaurre, S., Rodríguez, N., Quelopana, R., Valdivia, A., Rossi, E., & Nolasco-Mamani, M., 2024. Percepciones de la Inteligencia Artificial y su Impacto en la Integridad Académica en Estudiantes Universitarios de Perú y Chile: Una Aproximación a la Educación Sostenible. *Sustainability*. https://doi.org/10.3390/su16209005 .

Vinogradova, E. (2023). Artificial intelligence technologies and the rise of cyber threats in Latin America. Latinskaia Amerika, 3, 34. https://doi.org/10.31857/S0044748X0024415-5

Wack, P. (1985) 'Scenarios: Shooting the Rapids', *Harvard Business Review*, November-December, pp. 139-150.Weaver, T. (1971). The Delphi Forecasting Method. *The Phi Delta Kappan, 52*(5), 267–271. [in line] Available in: https://www.jstor.org/stable/i20372863 [Accessed 15 september 2024].

Winkler, J., y Moser, R., 2016. Sesgos en los estudios Delphi orientados al futuro: una perspectiva cognitiva. *Pronóstico tecnológico y cambio social* , 105, págs. 63-76. https://doi.org/10.1016/J.TECHFORE.2016.01.021 .

Zawacki-Richter, O., Marín, V. I., Bond, M. & Gouverneur, F. (2019) 'Systematic review of research on artificial intelligence applications in higher education – where are the educators?', International Journal of Educational Technology in Higher Education, 16, p. 39.

EFL Students’ Perspectives on ChatGPT in Translation: Exploring AI Assistance, Motivation, and Learning Outcomes

Cao-Tuong Dinh
English Language Department, FPT University, Can Tho Campus, Vietnam

TuongDC@fe.edu.vn

https://doi.org/10.34190/ejel.23.2.4006

Abstract: Recent advances in artificial intelligence (AI) offer promising opportunities to improve language education, particularly in translation, by providing tools that can enhance both learning processes and outcomes. Yet, how these AI tools are perceived and integrated, especially in areas that demand cultural sensitivity and a nuanced understanding, has not been fully explored, notably from the perspective of English as a Foreign Language (EFL) students. This study set out to examine how English major students view the use of ChatGPT, a text-based generative AI tool, within translation classes, using the Expectation-Value Theory as a framework. The study involved 62 junior English majors from a university in Vietnam and employed a mixed-methods approach, including pre- and post-course surveys, reflective journals, focus group interviews, and analysis of course grades. Results showed that students generally regarded ChatGPT as a helpful tool for improving translation accuracy, efficiency, and vocabulary skills. They valued its capacity to simplify complex translation challenges, improve sentence flow, and offer a variety of lexical choices, which in turn boosted their motivation and confidence. Students also mentioned that using ChatGPT helped promote collaborative learning by sparking more group discussions, which improved their translation skills. At the same time, they pointed out some limitations, especially how ChatGPT struggled with cultural nuances and idiomatic expressions. Because of this, students had to carefully review and adjust the AI’s suggestions themselves. The study points out that it’s really important to strike a balance between relying on AI tools and sticking with traditional, hands-on translation methods. Tools like ChatGPT can definitely support translation learning, but they can’t take the place of human judgment and effort. When looking at why students were motivated to use ChatGPT, the research found that it mostly came down to how helpful they believed the tool was, how much they valued using it, and how confident they felt about succeeding with it. These factors played a key role in their overall learning results. The study provides useful insights into how AI tools can support online learning by making it more efficient and engaging. However, it also reminds us that human judgment remains crucial, especially in translation tasks that involve cultural understanding. More research is needed to understand the long-term effects of AI in translation education and how well these tools work across different cultural settings.

Keywords: ChatGPT, AI in translation, Motivation, Engagement, Expectancy-value theory

1. Introduction

In recent years, artificial intelligence (AI) has made remarkable progress, reshaping many aspects of education, especially in language learning and translation. Among the various AI tools available, large language models (LLMs) have attracted growing interest because of their ability to support learners in a range of language-related activities, including generating text, answering questions, and translating (Barrot, 2024; Karataş, Yaşar, and Gunyel, 2024). Their capacity to produce text that closely resembles human writing has made them valuable additions to language and translation teaching, offering new possibilities that were difficult to achieve with traditional approaches (Barrot, 2024; Karataş, Yaşar and Gunyel, 2024). However, despite their growing use, the integration of AI into educational contexts, particularly in translation education, remains underexplored in empirical research (Sahari, Al-Kadi and Ali, 2023; Xiao and Zhi, 2024). The rapid growth of generative AI has prompted a need for more in-depth investigation into its applications, effectiveness, and challenges, especially when deployed as a learning tool in translation tasks (Lo, et al., 2024; Salloum, et al., 2024).

Recent studies have explored the implications of AI in second language acquisition (SLA) and translation pedagogy, highlighting both its advantages and limitations. AI tools have been shown to enhance students' engagement with linguistic content, offer personalized learning experiences, and provide immediate feedback that can foster self-regulated learning (Sahari, Al-Kadi and Ali, 2023; Xiao and Zhi, 2024). These tools also serve as cognitive scaffolds, assisting learners in tackling complex linguistic structures and increasing their autonomy in the learning process (Karataş, et al., 2024; Xiao and Zhi, 2024). However, despite these advantages, there remain substantial challenges that warrant further investigation. AI translation tools perform well with sentence structure and straightforward translations. However, they frequently struggle to capture deeper meanings, idiomatic expressions, and cultural nuances, areas where human judgment remains essential to ensure accuracy

Cite this article: Dinh, C-T. et al. 2025. “EFL Students’ Perspectives on ChatGPT in Translation: Exploring AI Assistance, Motivation, and Learning Outcomes”, *Electronic Journal of e-Learning*, 23(2), pp 99-116, https://doi.org/10.34190/ejel.23.2.4006

and quality (Ghassemiazghandi, 2024; Sahari, Al-Kadi and Ali, 2023; Van Horn, 2024). On top of that, the quality of AI translations can vary a lot depending on the language pair or the type of text, which makes it hard to rely on AI across all kinds of translation tasks. This shows that while AI has great potential, we also need research that looks beyond technology and considers its real-world limits, especially in teaching environments.

One major gap in current research is how students mentally engage with AI-generated translations, especially in advanced translation courses. While earlier studies have looked at AI's general usefulness for translation and language learning, few have explored how students critically interact with AI outputs or how these tools might actually help students develop deeper learning strategies in translation education. Additionally, the intersection between AI-assisted learning and self-regulated learning strategies, which are crucial for promoting autonomous learning, remains an area of limited research.

Focusing specifically on EFL learners is crucial because these students face unique linguistic and cultural challenges in translation tasks that AI tools often cannot fully address. Motivation plays a pivotal role in this context. Yet, there is limited emperical evidence on how students' motivation to use AI tools is influenced by their perceptions of these tools' utility, intrinsic value, and expectancy for success in translation activities. Understanding these motivational and cognitive dimensions is fundamental for efffectively integrating AI into translation pedagogy and designing instructional approaches that promote balanced reliance on technology and manual skills. Therefore, this study aims to fill these gaps by focusing on how English major students in advanced translation courses perceive the value of AI-driven tools, using the Expectation-Value Theory (Wigfield and Eccles, 2000) as the guiding framework. This theory underscores the role of students' motivation, shaped by their perceptions of AI's utility and relevance, in determining their engagement with AI tools and the effectiveness of these tools in enhancing their translation performance.

Despite these promising applications, AI tools have notable limitations, particularly when handling complex cultural nuances and idiomatic expressions. These challenges underscore the need for further investigation into the pedagogical effectiveness of AI tools in real-world translation tasks. While AI translation models are highly effective at processing syntax and literal translations, they frequently fall short in understanding context, tone, and cultural appropriateness, which are essential for high-quality translations (McIntosh, et al., 2025; Sahari, Al-Kadi and Ali, 2023). This research goes beyond simply evaluating the technical capabilities of AI in translation; it takes a closer look at its impact on teaching and learning. In particular, it focuses on how students perceive AI's role in improving translation accuracy and how well it fits with their learning goals. The study also considers the wider effects of bringing AI into the classroom, especially how these tools affect student motivation and learning outcomes. More specifically, it aims to understand how students' views on the reliability of AI and its potential to improve translation quality influence their overall engagement with these technologies. With AI becoming increasingly common in education, this study responds to the pressing need to explore how it can be integrated effectively into translation teaching while also recognizing its limitations and the importance of students critically interacting with AI outputs.

By examining EFL students' perceptions and mental engagement with AI in translation, this research hopes to offer useful insights on how AI can support learning in translation education, without losing sight of the crucial balance between technological support and human expertise.

2. Literature Review

2.1 ChatGPT's Role in EFL Learning and Language Education

Generative AI has been increasingly recognized as an important aid in language learning, offering capabilities such as text generation, question answering, and translation (Alawida, et al., 2023). Its role in language learning is increasingly being explored, with studies indicating its potential as a knowledgeable learning companion (Solak, 2024). Students have shown the capability of critically evaluating and modifying ChatGPT's outputs, which can enhance learning experiences and offset academic integrity concerns (Xiao and Zhi, 2023). Additionally, generative AI applications has been recognized for its usefulness in informal digital learning environments, where it supports EFL learners in engaging with language tasks creatively and productively (Liu and Ma, 2023; Li, 2024).

Several empirical studies underscore the pedagogical value of generative AI in enhancing learners' confidence and task completion, particularly in language-related assignments. Xiao and Zhi's (2024) study revealed that ChatGPT can facilitate students' tasks requiring language competence. In a similar vein, Van Horn (2024) looked at how Korean university EFL students feel about generative AI and found that many students had positive opinions. They especially liked how it helped improve their language accuracy and encouraged them to learn on their own. These findings show that generative AI has a lot of potential to make language learning better for EFL students, no matter which platform is used. In this study, Van Horn's results help us understand how EFL students might view generative AI in translation classes, particularly when it comes to how useful they find it and how it supports their independence as learners.

While generative AI has demonstrated significant value in improving language learning outcomes, further investigation is needed to understand how these tools are perceived by students in the context of advanced translation courses in Vietnam. Specifically, exploring students' engagement with AI tools, their critical evaluation of AI-generated outputs, and the impact of these tools on learner autonomy provides a clear research opportunity for future studies.

2.2 Using Generative AI in Translation

For instance, Ghassemiazghandi (2024) evaluated a large language model's translation accuracy using the BLEU score and found a notable improvement over traditional machine translation tools. Such findings affirm the growing accuracy and reliability of generative AI in performing translation tasks. Furthermore, studies in varied contexts, such as the Arab world, have found that AI-based tools are particularly effective in handling the mechanical aspects of translation, providing learners with a base version that can be reviewed and improved upon by humans (Ghassemiazghandi, 2024; Sahari, Al-Kadi and Ali, 2023). In informal digital settings, generative AI helps learners interact more with the material by allowing them to revise, reflect, and learn from AI-generated content right away. However, how people accept and view generative AI in translation isn't the same for everyone. Salloum and colleagues (2024) found that acceptance varies depending on the language pair, showing that AI tools need to be evaluated based on specific languages and contexts. There are also ongoing concerns about generative AI struggling with idioms, cultural details, and style. Because of these limitations, human review and editing remain important, especially for professional translations or those that require a deep understanding of context (Ruoqi, Yuan, and Gochuico, 2023; Sahari, Al-Kadi and Ali, 2023).In this study, such concerns are central to the investigation, as student reflections and evaluations provide insight into both the strengths and the limitations of using generative AI in translation education. This body of literature supports the importance of studying learners' cognitive and motivational engagement with AI outputs, especially when integrated into instructional settings.

While generative AI tools have demonstrated clear advantages in terms of efficiency and accuracy, there remains a need to investigate how these tools influence the learning outcomes of students in the context of translation education in both developed and emerging markets like Vietnam. This research can help determine the pedagogical strategies best suited for integrating AI tools into translation curricula, as well as how AI tools influence students' cognitive and motivational engagement in translation tasks.

2.3 Expectation-Value Theory in Language Learning

The Expectation-Value Theory (Wigfield and Eccles, 2000) suggests that students' motivation is influenced by two key factors: how much they believe they can succeed at a task (expectancies) and how important they see the task (task values). Expectancy beliefs refer to students' confidence in their ability to complete a task successfully. This confidence is shaped by their past experiences, their self-belief, and the current situation (Wigfield, Tonks, and Klauda, 2016). When students have a strong belief that they can succeed, they tend to put in more effort and stick with the task longer. Task values are multifaceted and typically decomposed into four distinct types (Wigfield, Tonks and Klauda, 2016; Trautwein, et al., 2012):

Attainment (Importance) Value: The significance an individual places on doing well on the task, often linked to identity and personal goals (Wigfield, Tonks and Klauda, 2016). For example, a learner who views translation competence as central to their identity would place high attainment value on translation tasks.

Intrinsic Value: The inherent enjoyment or interest a person finds in performing the task (Harackiewicz, Smith and Priniski, 2016). This reflects how engaging or pleasurable the learner perceives the task itself to be, independent of external rewards.

Utility Value: The perceived usefulness of the task for achieving future goals, such as career advancement or academic success (Canning and Harackiewicz, 2015). In the context of AI-assisted translation, utility value would reflect how learners perceive generative AI tools as beneficial for their learning or professional practice.

Cost (not requested but often included): The perceived negative aspects associated with task engagement, such as effort, time, or anxiety (Flake, et al., 2015).

These expectancy and value components jointly influence learners' academic outcomes, including their engagement, persistence, and performance (Gaspard, et al., 2015). For example, when learners expect to succeed and value the task highly, they are more likely to dedicate effort and achieve better results.

These expectancy and value components jointly influence learners' academic outcomes, including their engagement, persistence, and performance (Gaspard, et al., 2015). For example, when learners expect to succeed and value the task highly, they are more likely to dedicate effort and achieve better results.

This theoretical framework sheds light on decision-making, perseverance, and motivation by emphasizing how individuals evaluate the costs and benefits associated with their actions. Motivation tends to increase when individuals feel confident in their ability to overcome challenges using their skills and strategies, especially when they perceive that the rewards of achieving their goals outweigh the effort or resources required. The Expectation-Value Theory has been widely applied across various disciplines, such as in physical education (e.g., Shang, Moss and Chen, 2023), language learning (e.g., Sun, et al., 2023), and mathematics (e.g., Fong, et al., 2023).

This theory suggests that students' intention to use ChatGPT is influenced by their perception of its value (e.g., utility, intrinsic interest, or alignment with goals) as well as the perceived cost (e.g., time, effort, or ethical concerns), with both factors playing critical roles in shaping motivation and decision-making. The integration of artificial intelligence applications in educational settings is seen as a method to improve learning outcomes, with students' perceptions playing a crucial role in their behavioral intentions and actual use (Chan and Zhou, 2023; Sankaran, et al., 2023). Li (2024) used the Situated Expectancy-Value Theory to study what influences pre-service EFL teachers' willingness to use technology, emphasizing the role of both expectancy beliefs and task value in how technology is adopted. This is relevant to the current study, which looks at how EFL learners' motivation to use generative AI in translation is shaped by how useful they think the tool is, how well it fits their learning goals, and what results they expect.

This study provides a chance to explore how Expectancy-Value Theory can help us better understand how students use AI tools in translation education, especially in the context of Vietnam. By examining students' views on AI's usefulness and its connection to their goals, we can gain insight into the psychological reasons behind their willingness to use generative AI in translation tasks.

2.4 Integrating ChatGPT and Expectation-Value Theory

By integrating insights from studies on generative AI and the Expectation-Value Theory, we can better understand how EFL students view the use of ChatGPT in translation classes. Learners who regard ChatGPT as an effective and engaging tool (high utility and intrinsic value) and who believe in their ability to use it successfully (high expectancy) are likely to be more motivated and perform better in translation tasks (Lo, et al., 2024; Slamet, 2024).

For instance, if students find that ChatGPT provides accurate translations and quick feedback, they are likely to see the tool as very useful. If they also enjoy using ChatGPT and find it interesting, the intrinsic value for the tool increases. When students believe that ChatGPT will help them succeed in their translation work, their confidence in doing well goes up, which boosts their motivation (Hmoud, et al., 2024).

This study builds on these ideas to look at how motivation affects students' views and behavior when using AI in translation. The framework guides both the analysis of numbers and personal experiences gathered in the research. It helps explore how students' views of AI's usefulness and their confidence in success shape their learning overall. The study aims to show how AI tools like ChatGPT can be added to language courses to improve both student engagement and learning results.

Specifically, this research investigates how EFL students feel about using ChatGPT in translation classes and how it affects their motivation and translation skills. To get a full picture, the study uses a mixed-methods approach that combines both surveys and interviews. Specifically, the objectives are as follows:

- To investigate students' perceptions of ChatGPT's value in translation tasks.
- To explore students' comparisons between their own translations and those generated by ChatGPT.

To address these objectives, two research questions were proposed:

RQ1: How do EFL students perceive the application of ChatGPT in translation tasks?

RQ2: How do students compare their manual translations with those generated by ChatGPT, and what factors influence their evaluation?

In this study, the term factors refers to the main variables that emerged from survey responses, reflective journals, and focus group discussions. These include how useful students found the tool, how fluent the translations were, how well cultural aspects were handled, as well as students' motivation and mental involvement. These factors align with the study's theoretical framework and research design. Using a mixed-methods approach helps combine different types of data, allowing the study to include both measurable results and students' personal reflections in the analysis.

3. Method

3.1 Research Design

This study employed a convergent parallel mixed methods approach, which combines both quantitative and qualitative data to offer a more comprehensive understanding of the research issue. The quantitative data were collected using pre- and post-survey questionnaires and end-course exam grades, while qualitative data were gathered through reflective journals and focus group interviews. In this approach, the data from both strands were analyzed separately and then compared to cross-check and deepen the understanding of the research problem (Creswell and Creswell, 2018).

The use of a mixed-methods design is justified in this study for several reasons. First, quantitative methods, through surveys and exam scores, allow for the measurement of students' attitudes and the impact of the intervention on their academic performance. Meanwhile, qualitative methods, such as reflective journals and focus groups, enable the researcher to explore the students' personal insights, experiences, and reflections regarding the use of generative AI tools in translation tasks. The combination of both types of data provides a more holistic view of how AI-assisted learning impacts students' motivation and performance.

On the first day of the Translation 2 course, the researcher administered pre-survey questionnaires to both Class 1 (control group) and Class 2 (experimental group). Afterward, Class 2 used generative AI (e.g., ChatGPT) to generate translations of a given text, comparing the AI-generated translations with their own translations. In contrast, Class 1 followed the traditional method, with the teacher providing corrections and feedback on students' translations. A post-survey questionnaire was administered to both groups after they completed their end-of-course exams. The effectiveness of the intervention was assessed by comparing the two groups' end-course grades, analyzing the pre- and post-surveys, and evaluating students' perspectives through their reflective journals and focus group interviews conducted at the end of the course.

3.2 Participants

Sixty-two junior English majors from a private university in the Mekong Delta, Vietnam, participated in the study. The students were enrolled in two intact classes with the same curriculum, having studied "Translation 1" and now studying "Translation 2." The analysis of their pre-course translation performance, based on available grades, showed no significant differences between the two groups. Pre-course grades were analyzed using an independent samples t-test to assess any initial differences in performance between the two groups. The full statistical results can be found in the Findings section.

3.3 Intervention: Students' Using ChatGPT for Their in-class Translation Practice

Classes 1 and 2 were assigned to translate a given Vietnamese text into English in groups, using only paper and pen to prevent them from using the Internet for assistance. Some technical terms were provided to aid their work. Afterward, Class 2 was allowed to use the free version of ChatGPT to translate the same text into English in class. The teacher instructed them to compare their translations with ChatGPT's version and take notes on anything they learned from the comparison.

Meanwhile, Class 1 followed a different approach: the teacher randomly selected translations from 2–3 groups, analyzed them, and corrected mistakes with the entire class. Both classes met with the teacher twice a week, and the same procedures were repeated until the end of the course.

3.4 Research Instruments

This research employed pre-course and end-course grades (available on the university's website), survey questionnaires, reflective journals and focus group interviews. The questionnaire, adapted from Eccles and Wigfield (1995) and Zhu, et al. (2012), consisted of two sections: Section 1 surveyed students' demographic information, and Section 2 comprised an 11-item five-point Likert scale, with five items for expectancy beliefs measurement and six items for assessment of attainment, intrinsic, and utility values.

The next research tool was students' reflective journals or diaries. Reflective journals are particularly effective in capturing participants' time-related development, changes over time, and even potential causal relationships between variables (Dörnyei, 2007). This method encourages participants to document activities and reflections they consider meaningful, providing a rich source of qualitative data (Jacelon and Imperio, 2005). In this study, voluntary participants were asked to write self-reflective reports on their perceptions of the benefits and limitations of using ChatGPT to revise their groupwork translations. Guided questions such as "Do you think ChatGPT helps your translation version better for this week topic? How?" Or "Do you think ChatGPT makes your translation version worse? In what way it is worse?"

Participants were given the option to either write or audio-record their reflections on their computers and submit them through Google Forms, which were accessible exclusively to the researchers and the participants. They were encouraged to complete at least one reflective journal each week for a total of eight weeks. However, they had the flexibility to stop submitting reflections if they reported no new strategies to document.

The last research tool was focus-group interviews. Focus groups, involving participants who have experienced a specific situation, centralized the interaction among participants, rather than with the interviewer, with the interviewer serving as a moderator to facilitate discussion (Bryman, 2016). This dynamic can yield deep and insightful discussions (Dörnyei, 2007) and allow participants' views to emerge naturally (Cohen, Manion and Morrison, 2018). In this study, each group of Class 2 was invited to share their experiences of using ChatGPT for their translation tasks at their willingness. Eight out of ten groups agreed to participated in this phase of this data collection while the other two groups did not participate due to scheduling conflicts among their members.

The validity and reliability of the research instruments are supported through several measures outlined in the study. For the quantitative data, the internal validity of the questionnaire was ensured by (a) adapting items previously validated in earlier research (Eccles and Wigfield, 1995; Zhu, et al., 2012) and (b) conducting a piloting phase. Cronbach's Alpha values for the variables, all exceeding 0.7, indicated strong internal consistency (Table 2). Construct reliability was further demonstrated by consistent Cronbach's Alpha values across both the pre-survey and post-survey phases (Table 1). For qualitative data, thematic analysis conducted by the researcher and a colleague followed the guidelines established by Braun and Clarke (2006), bolstering the validity of the findings. An inter-rater agreement rate of at least 75% (Mackey and Gass, 2022) was used as a benchmark, and in this study, an 80% agreement was achieved. Any inconsistencies in coding were managed through deliberation or removal from the analysis.

3.5 Data collection and Analysis

3.5.1 Piloting phase

To ensure the internal consistency of the survey items and assess respondents' comprehension, a pilot survey was conducted with 35 junior English major students from the same research site. The participants received an email in Vietnamese detailing the research purpose, estimated survey duration, and a request for their voluntary consent to participate. Their responses were automatically recorded in Google Sheets, accessible exclusively to the researchers.

3.5.2 Data collection procedures for the official research

The official survey questionnaires were distributed to 62 participants via email, with email addresses obtained from the university's publicly accessible staff directory. To ensure participants' willingness to participate, an email was sent explaining the research purpose and including a link to the questionnaire along with a participant consent form. Participants had a right to discontinue participation at any time without any repercussions. Pre-survey data were collected from August 6 to 15, 2024, while post-survey data were gathered from October 15 to 22, 2024. The responses were automatically recorded in Google Sheets, accessible only to the researchers.

3.6 Reliability and Validity

The validity and reliability of the research instruments were ensured through several measures. For the quantitative data, the internal validity of the questionnaire was ensured by adapting items from previously validated instruments (Eccles and Wigfield, 1995; Zhu, et al., 2012) and conducting a piloting phase. Cronbach's Alpha values, indicating strong internal consistency, were above 0.7 for all variables.

Table 1: Construct reliability of the piloting phase

Variables	Cronbach's Alpha	N of items
Expectancy beliefs	.743	5
Attainment (Importance) value	.798	2
Intrinsic value	.707	2
Utility value	.725	2
Academic outcomes	.763	4

Table 2: Construct reliability of the actual phase

Variables	Cronbach's Alpha	N of Items
Expectancy beliefs	.716	5
Attainment (Importance) value	.815	2
Intrinsic value	.827	2
Utility value	.761	2
Academic outcomes	.908	4

4. Findings

4.1 RQ1: How do EFL students perceive the application of ChatGPT in translation tasks?

4.1.1 Quantitative Results

Pre- and post-tests

The mean of two groups' end-course grades was calculated to compare the differences in their translation performance. Results from the pre-course analysis showed no significant difference between the two groups, as detailed below in Table 3 and Table 4.

Table 3: Mean scores of the two groups before the treatment

Descriptive Statistics				
Class	N	Mean	Std. Deviation	Std. Error Mean
Class1	31	7.68	.86	.15
Class2	31	7.38	.71	.12

Table 4: Independent Samples t-Test Comparing Pre-course grades between groups

t	df	p	Mean diff.	SE Diff.	95% CI (Lower, Upper)
1.51	62	.136	.30	.20	-0.10, 0.70

Note. SE = standard error; CI = confidence interval

The next step was to compare the end-course exam grades to determine any differences in translation performance after the intervention. The post-course analysis showed that students in Class 2, who used ChatGPT as a tool for translation tasks, performed slightly better than those in Class 1, who followed the conventional approach. Table 5 below presents the comparison of the two groups' end-course exam grades.

Table 5: Comparison of the mean of the two groups' end-course exam grades

Group Statistics					
	Class	N	Mean	Std. Deviation	Std. Error Mean
Post-tests	Class1	31	8.04	.46	.08
	Class2	31	8.31	.58	.10

The mean end-course grades of Class 1 (control group) was 8.035 (SD = 0.4550), while Class 2 (experimental group) achieved a higher mean score of 8.310 (SD = 0.5776). This difference suggests that students in Class 2, who were allowed to use ChatGPT, performed slightly better in their translation tasks compared to those in Class 1B, who followed a conventional approach without AI assistance. However, an Independent Samples T-Test was performed to assure that this difference was statistically significant or not (Table 6).

Table 6: Independent Samples t-Test Comparing Post-tests Scores Between Groups

t	df	p	Mean diff.	SE Diff.	95% CI (Lower, Upper)
-2.08	62	.042	-0.27	0.13	-0.54, -0.01

Note. SE = standard error; CI = confidence interval. Levene's test was not significant (p = .080), so equal variances were assumed. Table 6 showed that the groups' scores were statistically significant. The t-test showed a statistically significant difference in the average scores of the two classes' post-tests (t = -2.076, df = 62, p = 0.042).

The mean difference of -0.27 indicates that Class 2 scored about 0.27 points higher on average than Class 1. The 95% confidence interval for this difference ranged from -0.54 to -0.01, suggesting that while the difference is small, it is statistically meaningful.

Additionally, the effectiveness of the treatment was also calculated by the experimental groups' end-course performance. Hence, the pre-test and post-test of the treatment group were compared (Table 7):

Table 7: Descriptive Statistics for Pre- and Post-Test Scores in the Experimental Group

Group Statistics					
	Tests	N	Mean	Std. Deviation	Std. Error Mean
Pre- and post-grades	Pre-test	31	7.38	.72	.13
	Post-test	31	8.31	.58	.10

To evaluate the impact of using ChatGPT, a paired-samples t-test was conducted to compare participants' scores before and after the intervention. The results are summarized in Table 8.

Table 8: Paired-Samples t-Test Comparing Pre- and Post-Test Scores

t	df	p	Mean diff.	SE Diff.	95% CI (Lower, Upper)
-5.59	62	< .001	-0.93	0.17	-1.26, -0.59

The comparison of pre- and post-test results reveals a significant improvement in the experimental group's performance after using ChatGPT. Table 8 presents the results of the independent samples t-test, which was conducted to assess the statistical significance of the difference between the two sets of scores.

The Levene's test confirmed that there was no significant difference in the variance between the pre- and post-test scores (p = 0.479), suggesting that the assumption of equal variances was not violated. The t-test results indicated a highly significant difference between the two sets of scores (p < 0.01). The mean post-test score was 8.310 (SD = 0.5776), which was higher than the pre-test score of 7.384 (SD = 0.7179), with a mean difference of -0.9258. This suggests that the experimental group's performance significantly improved after ChatGPT was incorporated into the translation process.

Pre- and Post-Surveys

To examine students' changes in their perceptions of using ChatGPT before and after the treatment, pre- and post-surveys were performed (Table 9). A paired t-test was performed to examine changes in students' perceptions of ChatGPT's effect on vocabulary and grammar before and after its use. This method was chosen because it accounts for the dependent nature of the data, as the same participants provided responses at two different time points (Field, 2018). The paired t-test is suitable for evaluating within-subject variations over time while controlling for individual differences (Pallant, 2020). Furthermore, it remains a valid statistical approach when the sample size is sufficiently large (typically n > 30-50), even if the data does not follow a normal distribution (Field, 2018).

Table 9: Paired-Samples t-Test Results for Pre- and Post-Surveys

Pair	t	df	p	Mean Diff.	SE	95% CI (Lower, Upper)
EB	-1.51	60	.136	-0.16	0.10	-0.37, 0.05
AV	-0.87	60	.387	-0.12	0.14	-0.41, 0.16
IV	-2.19	60	.033	-0.25	0.11	-0.47, -0.02
UV	-2.04	60	.046	-0.27	0.13	-0.54, -0.01
AA	-5.14	60	< .001	-0.67	0.13	-0.93, -0.41

The paired-samples t-test results provide a nuanced understanding of the changes in participants' scores between pre- and post-tests across different measures. While some variables demonstrated significant improvements, others remained unchanged.

Table 10 provides the descriptive statistics for participants' pre- and post-survey scores across five measured variables, showing the average values and variability at both time points.

Table 10: Descriptive Statistics for Pre- and Post-Survey Scores

Variable	Time point	*n*	Mean	SD	SE
EB	Pre	62	3.70	0.55	0.07
	Post	62	3.85	0.50	0.06
AV	Pre	62	3.91	0.84	0.11
	Post	62	4.03	0.69	0.09
IV	Pre	62	3.75	0.80	0.10
	Post	62	4.00	0.68	0.09
UV	Pre	62	3.78	0.79	0.10
	Post	62	4.05	0.74	0.09
AA	Pre	62	3.08	0.62	0.08
	Post	62	3.75	0.78	0.10

For instance, IV scores improved significantly from pre- to post-test ($t(60) = -2.19$, $p = 0.033$), with a small effect size (Cohen's $d = -0.28$). Similarly, UV scores showed a statistically significant increase ($t(60) = -2.04$, $p = 0.046$) with a small effect size (Cohen's $d = -0.26$). These findings suggest that participants experienced modest but measurable progress in these areas.

Table 11 presents the effect sizes for paired-samples comparisons, detailing the magnitude and confidence intervals of changes in each variable from pre- to post-test.

Table 11: Effect Sizes for Paired-Samples Comparisons

Pair	Cohen's *d*	95% CI (Lower, Upper)	Hedges' *g*	95% CI (Lower, Upper
EB	-0.19	-0.45, 0.06	-0.19	-0.44, 0.06
AV	-0.11	-0.36, 0.14	-0.11	-0.36, 0.14
IV	-0.28	-0.54, -0.02	-0.28	-0.53, -0.02
UV	-0.26	-0.52, -0.01	-0.26	-0.51, -0.01
AA	-0.66	-0.93, -0.38	-0.65	-0.93, -0.38

Note. Cohen's *d* and Hedges' *g* represent standardized effect sizes. Negative values reflect post-test improvements. CI = confidence interval; EB = Expectancy Beliefs, AV = Attainment Value, IV = Intrinsic Value, UV = Utility Value, AA = Academic Achievement.

The most substantial improvement was observed in AA scores, which demonstrated a highly significant increase ($t(60) = -5.14$, $p < 0.001$) and a moderate effect size (Cohen's $d = -0.66$), demonstrating a meaningful change and the strongest effect of the intervention among all variables. However, EB scores ($t(60) = -1.51$, $p = 0.136$) and AV scores (t(60) = -0.87, p = 0.387) showed no statistically significant differences, suggesting that the intervention did not substantially impact these measures.

These results highlight the targeted effectiveness of the intervention, particularly in improving AA scores, and to a lesser extent, IV and UV scores. The varying effect sizes indicate that while the intervention positively influenced some areas, its impact was not uniformly distributed across all measures.

4.2 Qualitative Results

4.2.1 Theme 1: Enhancing translation accuracy and vocabulary

Many participants perceived ChatGPT as a useful tool for the translation accuracy and helped them grow vocabulary through the use of contextually relevant word hints and insightful explanations for complexed or confusing epxressions. For instance, one student noted, “With ChatGPT, we can use the correct words for specific terms in the educational field, such as ‘học bạ’ as ‘school report’ and “tuyển thẳng” as “direct recruitment’” (Student 7, Diary 1). Similarly, one significant advantage of ChatGPT was its handling of colloquial idioms, such as “nét bình dị đậm chất miền Tây” as “the rustic charm of the Mekong Delta” Group2_Student 8). Another student valued the easy-to-understand explantions for technical terms so that they could make a right word choice for their context, noting, “It explains technical terms like the difference between ‘submerge’ and ‘immerse’” (Student 18, Diary 1).

Students also appreciated the forte of ChatGPT in providing new and suitable synonyms, as well as broad lexical repertoire. One student took notes in a journal, “ChatGPT replaced ‘emergency department’ with a more specific term such as “Emergency, Intensive Care, and Toxicology”, which helped improve clarity” (Student 4, Diary 2). Similarly, ChatGPT suggested a more accurate alternative like ‘psychological trauma’ instead of ‘traumatic events’ (Student 4, Diary 2). By offering contextual-based expressions, ChatGPT contributed to more natural translations (Student 9, Diary 1).

4.2.2 Theme 2: Enhancing efficiency and fluency

Students commended ChatGPT’s capacity to decrease the amount of time needed for revision while also enhancing the readability anf fluidity of translations. For instance, one participant reflected, “ChatGPT helps me get translations faster, which helps me to understand the content better” (Student 15, Diary 4), while another added, “It processes and translates long sentences efficiently, and so I can save time for the task” (Group2_Student 7).

Beyond providing rapid suggestions for translated text, ChatGPT’s translations can help enrich sentence structure and reader-friendliness. A student shared, “ChatGPT transforms complex sentences into simpler and more coherent ones; this not only helps enhance readability but also improve sentence coherence” (Student 10, Diary 2). Additionally, it helped students avoid run-on sentences. “Its suggestions help reduce run-on sentences and improve sentence flow”, shared Student 9 (Group2).

4.2.3 Theme 3: Enriching learning and collaboration

ChatGPT plays a role of a learning aid when expanding students’ vocabulary and improving their sentence structures. Many participants found that ChatGPT supported their learning through topic-specific vocabulary. For instance, one student noted, “I learned precise terms like ‘fields of study’ instead of ‘majors’ in the phrase “Hanoi Pedagogical University 2 is recruiting about 2,000 students for 23 majors“ and applied them to other contexts such as “Admission selection for fields such as Preschool Education, Physical Education” (Student 9, Diary 1). Another favored better terms that ChatGPT suggested for tourism topics, such as the expression of ‘beautiful place’ substituted by ‘tourist spot’ in “This tourist spot is beautiful thanks to nature and its altitude advantage. Here “điểm du lịch” literally means “tourist spot,” a more precise term than just “beautiful place” (Student 12, Diary 2).

Furthermore, ChatGPT also facilitated group discussions and enhanced collaborative analysis. By comparing manual translations with ChatGPT’s version and their manual work, students could identify errors and make improvements. One participant shared, “ChatGPT’s translations help us analyze and understand different styles during group discussions. For example, one group’s trasnslation of “Rừng già xếp tầng lớp dưới những khối mây có cảm giác như đất trời hòa làm một” was “The ancient forest is layered under clouds, which gives a feeling that earth and sky merge as one” which, I think, is better than ours “The ancient forest grows layers by layers under clouds which makes us have the feeling that earth and sky merge are one.” (Group1_Student 4).

4.2.4 Theme 4: Concerns about dependency

Despite its strengths, students were concerned about ChatGPT's influence on their possibility of being over-reliant on ChatGPT, which could thwart their critical thinking and empede their effort to unpack meanings from contexts. Some expressed this fear, "ChatGPT makes it easy to depend on its outputs, which can prevent me from developing my own skills" (Student 6, Diary 4). Another echoed this concern and added "Over-reliance on ChatGPT reduces the learning I gain from manual translations, sometimes I feel less confident about my [translated] version since I fear that my word choice or grammar is not academic enough" (Group2_Student 8).

Beyond concerns about over-reliance on ChatGPT, students also faced other considerations, such as handling cultural and contextual nuances in the terms they used during their learning and translation techniques. For example, a participant pointed out, "ChatGPT struggles with cultural terms like "công đất" translating them awkwardly as "5 công of land" (Group1_Student 5), or "miền Tây" translated as "the Western" (Group 1_Student 6). Similarly, for idiomatic phrases, ChatGPT often missed the tone, requiring manual adjustments: "For idiomatic phrases, ChatGPT often misses the tone, requiring manual adjustments" (Student 13, Diary 4).

In addition to the themes identified from both the reflective journals and focus group interviews, additional themes emerged uniquely from each data collection method. Specifically:

4.2.5 Theme 5: Increasing confidence in translation (from reflective journals)

Although some students raised concerns about feeling less confidence in their manually translated version, most participants admitted a significant benefit of ChatGPT is its ability to enhance the quality of their translations for complex translation texts. By simplifying confusing and challenging phrases and terminologies, ChatGPT provided them with a more natural and contextual-bound terms, hence increased their confidence on the submitted tasks. For instance, a student was satisfied with the tool for its precise vocabulary for technical terms like 'psychological trauma,' 'toxicology,' and 'dermal injury,' which they had not encountered before (Student 4, Diary 2). Similarly, another student appreciated ChatGPT's ability to accurately translate specific terms such as 'post-traumatic stress disorder' as 'rối loạn căng thẳng sau chấn thương (PTSD)' and 'mechanical ventilation' as 'thở máy,' which enhanced their understanding and precision in medical translation (Student 9, Diary 2). This capability of ChatGPT not only improves the accuracy of targeted language translations but also encourages students to approach unfamiliar topics with greater confidence.

In addition to providing accurate terminology, ChatGPT helps reduce the anxiety often associated with tackling complex sentence structures, with several students acknowledging that using the tool eased their stress during translation tasks. For instance, one student shared, "Using ChatGPT makes me feel less overwhelmed when I have to turn long Vietnamese sentences into proper English, especially when I'm unsure how to arrange the ideas" (Student 13, Diary 4). Another student appreciated how the tool quickly breaks down complicated sentences into clear, manageable parts, helping them save time and reduce frustration. This allowed them to concentrate more on making their translations accurate (Student 15, Diary 4). By easing these difficulties, ChatGPT creates a more encouraging learning environment and boosts students' confidence as they improve their translation skills.

4.2.6 Theme 6: Limitations in translation quality (from reflective journals)

While ChatGPT offers several advantages, interestingly, some students also identified significant limitations in its translation quality, particularly regarding the naturalness and accuracy of its outputs, especially when the terms related to cultural aspects. For example, a student noted that *"ChatGPT's version 'The ancient forest is layered under clouds, which gives a feeling that earth and sky merge as one' sounds better than ours, but still feels a bit stiff and misses the poetic feel of the original Vietnamese description"* (Group1_Student 4). Similarly, another student reflected that sometimes ChatGPT's suggestions for overly academic or high-level vocabulary may not align with the intended tone of the translated text. For example, one diary noted "ChatGPT's word choices like 'altitude' and 'dense canopy' sound impressive and fancy, but don't always fit the relaxed and inviting mood of a travel description. This mismatch can make the translation feel less reader-friendly" (Student 17, Diary 4). These limitations suggest that while ChatGPT is effective for technical accuracy, it may lack the adaptability and cultural sensitivity required for more nuanced and contextually appropiate text-type style translations.

In addition to sounding overly formal, students reported instances of mistranslations and contextual errors that required manual intervention. For instance, one student remarked that "ChatGPT sometimes uses weird or incorrect words for the context," which can confuse the intended meaning (Student 5, Diary 2). For instance, the homestay owner's activities like *"bắt sò huyết"* (catching cockles) or *"bơi xuồng dỡ lú"* (checking fish traps by boat) reflect local cultural aspects that ChatGPT might not render the naturalness or accuracy of the terms without intervention. Such mistranslations risk losing the nuance and authenticity crucial for effective communication, especially in community-based tourism contexts. Another noted that "some expressions don't sound natural and require manual adjustments," such as the phrase *"Cung đường bám theo lưng chừng núi"* (translated as "The route follows the mountainside"), which may sound technically correct but can lack natural fluency and vividness, hence needing human rephrasing for a more engaging and clear description. These findings emphasize the importance of critical engagement with ChatGPT's responses and require students' cultural knowledge and ability to refine its translations to align with the tone and context of the targeted translated text. While acknowledging the undeniable value of ChatGPT's support in text translation, translators should consider the aforementioned limitations to maintain a balance between AI assistance and manual translation efforts.

4.2.7 Theme 7: Facilitation of collaborative learning (from focus group interviews)

Many students from the focus group interviews said that owing to the suggested version from ChatGPT, their groups worked together again, compared their manual version, and then sharpened their final version before offcial submission to the teacher. These times when they felt most exploited from the ChatGPT to enrich knowledge as well as boosted teamwork. One participant shared, "We analyzed ChatGPT's translated outputs together to see whether they are correct, approriate or not," (Group1_St4). Similarly, several emphasized that by comparing ChatGPT's with their manual translations, they developed the ability to identify errors and engaged in critical discussion to have a better version of their final product (Group 2_St1). This process not only enhances students' translation skills but also promotes teamwork, as group members actively contribute their own perspectives and insights to refine their collective understanding of the required translation tasks.

In addition to strengthening collaborative analysis, ChatGPT aided thematic vocabulary building, hence further enriched the vocabulary learning experience. Students frequently used the tool to find out and document new vocabulary relevant to specific topics, which they afterwards incorporated into exams and other assignments. One student commented, "We noted down useful words by themes and applied them in our tests as well," indicating that ChatGPT can facilitate targeted vocabulary acquisition (Group1_St2). By assisiting students in categorizing and retaining new vocabulary items systematically, ChatGPT enhances their lexical repertoire and prepares them for future translation tasks. Together, using ChatGPT in groupwork for translation tasks creates an interactive and dynamic learning environment where students harness ChatGPT's capabilities to optimize not only their individual performance but also their collaborative outcomes.

Figure 1 below illustrates these themes more vividly:

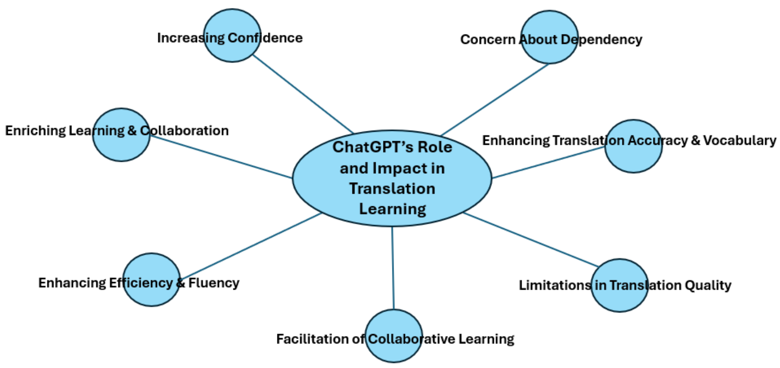

Figure 1: Students' perceptions of the application of ChatGPT in translation learning

4.3 RQ2: How do Students Compare Their Manual Translations With Those Generated by ChatGPT, and What Factors Influence Their Evaluation?

4.3.1 Theme 1: Comparative advantage of ChatGPT translations

Students consistently highlighted the positive side of ChatGPT since it can produce fluent, natural, and polished translations that often surpassed their manual ones in certain contexts. One student remarked, "ChatGPT is normally able to suggest academic words, and so its translation is with high level of naturalness and smoothness, which is, I think, closer to authentic papers" (Student 10, Diary 1). Similarly, another pointed out, "I like ChatGPT because it helps with difficult phrases and makes the translations more concise and formal" (Student 8, Diary 4). In group discussions, participants emphasized the superior fluency of ChatGPT's outputs, "ChatGPT's translations feel smoother and more polished than our manual versions", one student acknowledged (Group2_St10).

Another key advantage was the superiority of ChatGPT's provision of the variety of synonyms and diverse structures compared to students' manual outcomes. ChatGPT was able to provide alternative patterns and synonyms relevant to specific technical contexts, which helped facilitate the refinement of their self-translated versions. One participant shared, "ChatGPT gives various word options and diverse structures that we can consider the one that is most appropriate for the context" (Student 7, Diary 1). Another student appreciated ChatGPT's ability to offer appropriate technical terms such as 'psychological trauma' instead of 'traumatic events,' indicating that it could demonstrate a deeper, more specific contextual understanding (Student 4, Diary 2). This ChatGPT's capacity allowed students to explore and select the most suitable expressions to enhance the accuracy and contextual relevance of their translations.

4.3.2 Theme 2: Preference for manual translations in specific contexts

Despite ChatGPT's advantages, students often preferred their manual translations in scenarios requiring cultural sensitivity or personal styles. Many expressed those manual efforts better captured nuanced meanings and cultural references. For example, one participant reported, "I sometimes see sentences [suggested by ChatGPT] that don't sound better than my own ones" (Student 18, Diary 2). Another added, "ChatGPT doesn't fully convey cultural nuances, especially idioms or specific topics related to local dialects" (Student 15, Group 4). This limitation was evident in examples like the mistranslation of '05 công đất' as '5 công of land,' which manual version can be much more accurate (Group2_St9). Beyond limitations in obtaining cultural accuracy, several students found that they felt satified with their manual translations. One participant shared, "There are sentences where we feel our translations are better, so we don't use ChatGPT's version" (Student 4, Diary 3). Another remarked, "Given that AI models occasionally provide incorrect or misleading answers and tend to obscure my uncertainties, I am likely to favor my own translations", (Student 1_Diary 4). These insights indicate that although ChatGPT are superior to students' manual translation in terms of technical and formal texts, manual translations remain crucial for capturing cultural and personal nuances.

4.3.3 Theme 3: ChatGPT as a complementary learning tool

Alongside serving as a translation-assisted tool, the particpants regarded them as a valuable resource for refining and improving their manual translations. Many participants emphasized its role in enhancing readability and sentence smoothness. For example, one student shared, "ChatGPT is helpful for fine-tuning translations, especially when I need more comprehensive expressions" (Student 10, Diary 2). Several are inclined to use it for improving and varying sentence structures, stating, "It helps improve the coherence of our translated versions by suggesting better structures and transition words" (Student 13, Diary 3). Additionally, ChatGPT was also useful for stylistic learning, such as providing variations in tone and phrasing. One participant said, "We asked ChatGPT for more formal expressions and described the context in which the sentence appears, then learned from its suggestions" (Group2_St7).

Students characterized ChatGPT as a comparative learning tool that helped them identify and correct errors in their manual translations. One participant explained, "ChatGPT helps identify mistakes in our translations, such as typos, subject-verb agreement errors, or even the logical of thought. Thanks to this, we can fix them before submitting our paper to the teacher" (Group2_St7). These reflections accentuate ChatGPT's potential as a complementary resource that supports students in fine-tuning their written products while promoting critical engagement with their work.

4.3.4 Theme 4: Factors influencing translation evaluation

Some students acknowledged their shortcomings in advanced grammar and organizing ideas, and noted that using ChatGPT could help compensate for these weaknesses. ChatGPT has the cappability of producing grammatically and semantically coherent and cohesive sentences, which is conducive to students' learning of grammar and idea organization. One participant shared, "ChatGPT's alternative suggestions for our long sentences are helpful since they don't contain grammatical errors like run-on sentences and promote concise sentences that clearly convey multiple messages" (Group2_St9). Additionally, ChatGPT's suggestions were often appreciated for their fluency and alignment with intended meanings. A participant stated, "ChatGPT translates more naturally and more closely in meaning to the targeted text than our versions, and even more so than Google Translate's" (Group1_St2).

However, regarding unpacking meanings from some particular topics, especially Vienamese language when dialects exist among different regions, students indicated ChatGPT's limitations in translating this aspect. One student said, "I tried translating phrases like "bắt sò huyết" and "nhổ bồn bồn," but ChatGPT couldn't capture the local nuances; it translated them literally, which might confuse international readers' (Student 1, Group 8)." These evaluations indicate that while ChatGPT performs well in structured contexts, its effectiveness diminishes in informal or culturally nuanced scenarios, necessitating manual adjustments.

4.3.5 Theme 5: Balancing AI and human effort

Appreciating ChatGPT's benefits and recognizing its limitations, many students expressed the importance of individual effort in text-based translation for personal development and the end-course exam, while emphasizing the value of integrating ChatGPT for consultation purposes. Many described ChatGPT as a valuable learning and refinement tool that enhanced the accuracy and clarity of their translations. For example, one student shared, "I usually ask ChatGPT to check my translated text; and it normally provides some key improvements that helps me refine my sentence" (Student 15, Diary 2). This reflects how students integrate ChatGPT into their learning process, using it as a complementary tool rather than a replacement one.

At the same time, students emphasized the importance of individuals' critical engagement and self-reliance in their translation process. Several students said that they normally do not use the ChatGPT's version without discussing in groups for the final version. One participant said, "We follow three steps: at first, we still translate manually; then we ask ChatGPT for help; and finally, we check ChatGPT's suggestions with teammates. By doing so, we can learn better" (Student 2, Group 7). Others prefered the opposite way, "whenever applicable, we put and paste the text to ChatGPT.com, then discuss its suggestion in groups and decide what to keep and what to remove from the sentence" (Student 1, Group 5). This mindset ensures that students remain actively engaged in the translation process while leveraging ChatGPT's capabilities to enhance their work.

4.3.6 Theme 6: Balancing ChatGPT's strengths with student autonomy

Although the tempatation to fully using ChatGPT in their learning was huge, some students expressed concerns over its influence on their their learning autonomy and learning outcomes. One participant said, "If I just copy entire sentences or passages into ChatGPT, I'll gradually lose my ability to translate" (Student 3_Group3). Others highlighted how important it is to balance manual work with AI-generated results. For example, a student noted, "We should only use ChatGPT as a reference tool to check clarity, make more concrete sentences, and not rely on it entirely" (Student 4_Group1). Another emphasized the need to balance ChatGPT's use with their own critical thinking, "If I rely on ChatGPT for everything without making an effort to translate manually, my [translation] skills will erode, which is definitely detrimental to the final exam and my future as well" (Student2_Group 7).

In summary, while ChatGPT is perceived as a valuable tool for improving translation accuracy, fluency, and vocabulary, students are also aware of its limitations, particularly in handling cultural nuances and idiomatic expressions. The tool can significantly enhance translation efficiency, support collaborative learning, and boost student confidence. However, concerns about dependency and the need for manual evaluation remain prominent. Ultimately, the findings suggest that ChatGPT is most beneficial when used as a supplementary tool that supports, rather than replaces, the manual translation process, and is not a one-size-fits-all solution. This balanced approach allows students to fine-tune their translations while retaining critical thinking skills, ensuring that they develop both their linguistic and analytical abilities.

Figure 2 below provides a clear illustration of these themes.

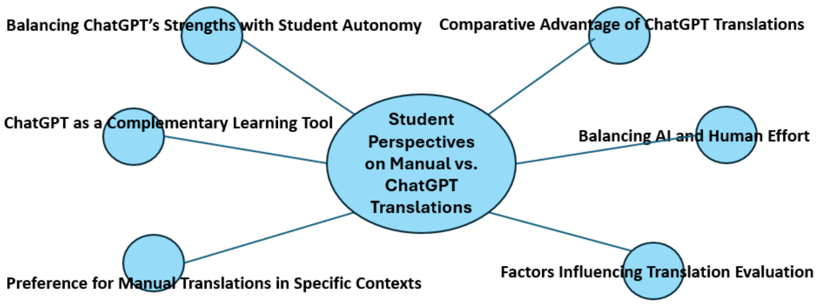

Figure 2: Student Perspectives on Manual vs. ChatGPT Translations

5. Discussion

While the findings confirm the positive effect of integrating ChatGPT into translation tasks on students' translation performance, the overall impact of these improvements is modest. This raises thought-provoking questions regarding the practical ramifications of AI tools in language learning. For example, the small mean difference between pre- and post-test scores (a 0.27-point increase in the experimental group compared to the control one) suggests that factors such as task complexity, students' baseline competencies, and the degree of human evaluation or judgement might moderate the Generative AI tool's effectiveness. This aligns with the work of Sahari, Al-Kadi and Ali (2023), who found that while AI tools like ChatGPT can aid in mechanical translation tasks, their impact on overall learning outcomes can vary significantly across individuals and context-bound settings.

A critical discussion of the quantitative results reveals that although ChatGPT improves fluency and accuracy, its impact may be constrained by its inability to fully capture cultural nuances and informal language styles. This limitation is significant because, as noted by Sahari, Al-Kadi and Ali (2023), cultural sensitivity is crucial for high-quality translation. Therefore, while ChatGPT can act as an efficient support tool, educators must ensure that it is complemented by traditional teaching methods that develop students' critical thinking and contextual analysis skills.

Furthermore, the qualitative findings emphasize the roles of ChatGPT in fostering teamwork's learning collaboration and enriching students' understanding of translated texts. Students reported that comparing their manual translations with those generated by ChatGPT stimulated critical thinking and group discussions, which lends support to the findings by Ghassemiazghandi (2024) on the collaborative potential of AI in educational settings. This highlights that while ChatGPT can promote individual learning autonomy (Xiao and Zhi, 2023), it can also amplify benefits in group activities. The sharing and reflections from focus groups and reflective journals suggest that the foremost benefit of ChatGPT resides in its capacity to boost collaborative learning, especially in translation tasks where the translated outcomes require not only grammatical and semantic accuracy but also culturally embedded meaning. When students engage in discussions comparing AI outputs with their own work, they not only identify mechanical errors and issues related to meaning but also support each other in meaning negotiation.

While Van Horn's (2024) study emphasized ChatGPT's role in improving general language skills among Korean university students, this study provides specific insights into how ChatGPT supports task-specific learning, particularly translation. While Van Horn reported students' positive attitudes towards the potential of AI tools like ChatGPT, our findings delve deeper into students' reflective processes, supported with focus group interviews, showing how they integrate ChatGPT's suggestions with their manual efforts for better learning outcomes.

Furthermore, this study is consistent with a study by Sahari, Al-Kadi and Ali (2023) that highlighted ChatGPT's advantages in mechanical translation processes while admitting its limitations in processing cultural aspects and idiomatic expressions. However, our study advances this discourse by illustrating how students leveraged the tool as a supporting learning aid, enriching their understanding through group discussions, a dimension not previously explored in the previous study.

Unlike earlier studies that primarily used observational or survey methods (e.g., Sahari, Al-Kadi and Ali, 2023; Salloum, et al., 2024), this research employed a convergent parallel mixed-methods design, combining pre- and post-surveys, reflective journals, and focus group interviews. This methodological triangulation provided deeper insights into how students integrate ChatGPT into their translation practices and the resulting impact on their skills and perceptions. For instance, the significant improvement in students' post-test scores, alongside qualitative feedback, indicates that ChatGPT can enhance learning outcomes when used as a complementary tool rather than a standalone solution.

Additionally, the use of the Expectation-Value Theory (Wigfield & Eccles, 2000) to interpret students' perceptions offers a novel perspective. The findings reveal that students' motivation to use ChatGPT was influenced by its perceived utility, intrinsic value, and their expectancy for success. For instance, participants valued ChatGPT for its ability to simplify complex terms and reduce anxiety around translation tasks, which aligns with the theory's emphasis on task value and expectancy beliefs as motivators. This theoretical framing distinguishes the study from prior research, providing a more structured understanding of students' attitudes and behaviors.

5.1 Contribution of the Current Study to the Existing Body of Literature

This study contributes to the growing body of research on the role of AI tools in language learning, particularly in translation education. By investigating EFL students' perspectives on using ChatGPT, the study expands upon existing findings and provides important insights into how AI tools can enhance learning outcomes and student engagement in translation tasks.

In line with prior studies (Xiao and Zhi, 2024; Van Horn, 2024), which emphasized the potential of AI in improving language accuracy and fostering autonomous learning, this research adds depth by exploring the intersection of AI use with self-regulated learning strategies. It builds on the work of Sahari, Al-Kadi and Ali (2023), who highlighted the benefits of AI tools for improving the mechanical aspects of translation (e.g., syntax and fluency). This study further confirms these findings while also emphasizing AI's limitations in addressing cultural nuances and idiomatic expressions, highlighting the need for human oversight in complex translation tasks.

Moreover, this study provides valuable insights for e-learning practices by showcasing how ChatGPT can be integrated into translation courses to enhance collaborative learning and critical thinking. Students found that using ChatGPT encouraged them to engage critically with the content, which resonates with generative AI's role in scaffolding learning (Karataş, et al., 2024). This collaborative element is crucial for the design of effective e-learning environments, where AI serves as a supporting tool rather than a replacement for human interaction.

In practical terms, this research offers useful recommendations for educators considering AI tools in their curriculum. ChatGPT, as this study shows, can significantly reduce time spent on routine translation tasks, allowing students to focus more on content comprehension and higher-order translation skills. However, educators must ensure that AI is used in a way that fosters student autonomy and encourages critical engagement, thereby preventing over-reliance on AI outputs. This insight can inform curriculum design and assessment strategies in e-learning environments, ensuring that AI tools complement traditional learning methods rather than overshadow them.

6. Conclusion

This study investigated English majors' perspectives on using text-based generative AI (specifically ChatGPT) in translation classes, employing the Expectation-Value Theory (Wigfield and Eccles, 2000) as a framework for understanding its impact on students' motivation, learning outcomes, and translation practices. The findings revealed that students valued ChatGPT as a valuable tool for providing timely and quality feedback on translation accuracy, efficiency, and vocabulary enrichment. In addition, ChatGPT plays a significant role in boosting students' motivation and confidence, and supporting their self-regulation and autonomous learning.

Despite these advantages, the study also identified critical limitations of ChatGPT, particularly in handling cultural nuances and idiomatic expressions. In these cases, students revised the suggested translated texts based on their teamwork discussions to accurately convey the meaning of both the source and target languages. A typical example of this is the translation of "05 công đất" or "miền Tây." These limitations highlight the importance of human review and critical engagement with AI-generated content. Therefore, while ChatGPT can mostly enhance translation fluency and accuracy, it should be considered and integrated as a complementary tool rather than a replacement for human expertise.

6.1 Strengths and Weaknesses of the Study

A major advantage of this study was its mixed-methods design, which combined a survey questuionnaire, pre- and post-tests from end-course exams, reflective journals, and focus group interviews to provide a comprehensive understanding of students' perceptions and experiences in utilizing ChatGPT during their translation classes. The pre- and post-tests objectively measured students' translation performance, while reflective journals and focus group interviews offered valuable insights into their learning processes. Additionally, the use of the Expectancy-Value Theory allowed for a structured analysis of students' motivation, contributing to the theoretical understanding of Generative-AI tools integration in higher education settings, particularly in translation skills practice.

However, certain limitations should be acknowledged. The study focused exclusively on English majors in Vietnam, which may limit the generalizability of the findings to other linguistic and cultural contexts. Additionally, the study's short-term intervention, lasting eight weeks and relying on pre- and post-tests, did not capture the long-term effects of AI-assisted translation learning.

6.2 Future Directions for Research

Future research should focus on several key areas to expand our understanding of AI tools in translation education. First, due to the current study's limitation in short-term intervention, longitudinal studies could investigate the long-term effects of AI on translation skills and student motivation. Second, cross-cultural comparisons could broaden the scope of research by including diverse linguistic and cultural contexts. For example, research that includes students from different backgrounds would provide a more comprehensive understanding of AI's effectiveness in various educational settings. Lastly, this study highlights the important role of collaborative learning among team members but does not explore the potential of peer feedback from other groups. Therefore, future research could further investigate this issue to better understand social learning dynamics in translation education.

AI Statement: The authors declare that Artificial Intelligence tools were not used in this study for the paper's conception, revision, or the creation of figures and tables.

Ethics Statement: Prior to the data collection, ethical approvals have been obtained from all participants.

References

Alawida, M., Mejri, S., Mehmood, A., Chikhaoui, B. and Isaac Abiodun, O., 2023. A comprehensive study of ChatGPT: advancements, limitations, and ethical considerations in natural language processing and cybersecurity. *Information*, [e-journal] 14(8), pp.1 - 23. Available at: <https://www.mdpi.com/2078-2489/14/8/462> [Accessed 03 June 2024].

Barrot, J., 2024. ChatGPT as a language learning tool: An emerging technology report. *Technology, Knowledge and Learning*, [e-journal] 29(2), pp.1151 - 1156. https://doi.org/10.1007/s10758-023-09711-4.

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, [e-journal] 3(2), pp.77 - 101. https://doi.org/10.1191/1478088706qp063oa.

Bryman, A., 2016. *Social research methods*. 5th ed. Oxford: Oxford University Press.

Canning, E.A. and Harackiewicz, J.M., 2015. Teach it, don't preach it: The differential effects of directly-communicated and self-generated utility–value information. *Motivation science*, *1*(1), 47. https://doi.org/10.1037/mot0000015

Chan, C.K.Y. and Zhou, W., 2023. An expectancy value theory (EVT) based instrument for measuring student perceptions of generative AI. *Smart Learning Environments*, [e-journal] 10(1), pp.1 - 22. https://doi.org/10.1186/s40561-023-00284-4.

Cohen, L., Manion, L. and Morrison, K., 2018. *Research methods in education.* 8th ed. New York: Routledge.

Creswell, J.W. and Creswell, J.D., 2018. *Research design: Qualitative, quantitative, and mixed methods approaches*. 5th ed. London: Sage.

Dörnyei, Z., 2007. *Research methods in applied linguistics*. Oxford: Oxford University Press.

Eccles, J.S. and Wigfield, A., 1995. In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, [e-journal] 21(3), pp.215 - 225. https://doi.org/10.1177/0146167295213003.

Field, A., 2018. *Discovering statistics using IBM SPSS statistics*. 5th ed. California: SAGE Publications.

Flake, J.K., Barron, K.E., Hulleman, C., McCoach, B.D. and Welsh, M.E., 2015. Measuring cost: The forgotten component of expectancy-value theory. *Contemporary educational psychology*, [e-journal] 41, pp.232-244. https://doi.org/10.1016/j.cedpsych.2015.03.002

Fong, C.J., Lee, J., Krou, M.R., Hoff, M.A., Johnston-Ashton, K., Gonzales, C. and Beretvas, S.N., 2023. Meta-analyzing the factor structure of the learning and study strategies inventory. *Journal of Experimental Education*, [e-journal] 91(2), pp.380 - 400. https://doi.org/10.1080/00220973.2021.2021842.

Gaspard, H., Dicke, A.L., Flunger, B., Brisson, B.M., Häfner, I., Nagengast, B. and Trautwein, U., 2015. Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental psychology*, [e-journal] 51(9), pp.1226 - 1240. https://doi.org/10.1037/dev0000028

Ghassemiazghandi, M., 2024. An evaluation of ChatGPT's translation accuracy using BLEU score. *Theory and Practice in Language Studies*, [e-journal] 14(4), pp.985 - 994. https://doi.org/10.17507/tpls.1404.07

Harackiewicz, J.M., Smith, J.L. and Priniski, S.J., 2016. Interest matters: The importance of promoting interest in education. *Policy insights from the behavioral and brain sciences*, *3*(2), pp.220 - 227. https://doi.org/10.1177/2372732216655542

Hmoud, M., Swaity, H., Hamad, N., Karram, O. and Daher, W., 2024. Higher education students' task motivation in the generative artificial intelligence context: The case of chatgpt. *Information*, [e-journal] 15(1), pp.1 - 18. Available at: <https://pdfs.semanticscholar.org/f010/9696f537e6d59070bdeb9594ceac38925b70.pdf> [Accessed 03 June 2024]

Jacelon, C.S. and Imperio, K., 2005. Participant diaries as a source of data in research with older adults. *Qualitative health research*, [e-journal] 15(7), pp.991 - 997. https://doi.org/10.1177/1049732305278603.

Karataş, F., Yaşar, F. and Gunyel, F.O., 2024. Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, [e-journal] 29(15), pp.19343 - 19366. https://doi.org/10.1007/s10639-024-12574-6.

Li, S., 2024. Exploring pre-service foreign language teachers' intention to adopt technology: An investigation from the perspective of situated expectancy-value theory. *Journal of China Computer-Assisted Language Learning* [e-journal]. https://doi.org/10.1515/jccall-20240012.

Liu, G.L. and Ma, C., 2023. Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, [e-journal] 18(2), pp.125 - 138. https://doi.org/10.1080/17501229.2023.2240316.

Lo, C.K., Yu, P.L.H., Xu, S., Ng, D.T.K. and Jong, M.S.Y., 2024. Exploring the application of ChatGPT in ESL/ EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, [e-journal] 11(1), pp.1 - 24. https://doi.org/10.1186/s40561-024-00342-5.

Mackey, A. and Gass, S.M., 2022. *Second language research: Methodology and design*. 3rd ed. New York: Routledge.

McIntosh, T.R., Susnjak, T., Arachchilage, N., Liu, T., Xu, D., Watters, P. and Halgamuge, M.N., 2025. Inadequacies of large language model benchmarks in the era of generative artificial intelligence. *IEEE Transactions on Artificial Intelligence*. https://doi.org/10.1109/TAI.2025.3569516

Pallant, J., 2020. *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS*. 7th ed. New York: Routledge.

Ruoqi, S., Yuan, X. and Gochuico, M.T., 2023. Translators and their use of ChatGPT. *Academia Lasalliana Journal of Education and Humanities*, [e-journal] 5(1), pp.49 - 59. https://doi.org/10.55902/PLRU9197.

Sahari, Y., Al-Kadi, A.M.T. and Ali, J. K. M., 2023. A cross sectional study of ChatGPT in translation: Magnitude of use, attitudes, and uncertainties. *Journal of Psycholinguistic Research*, [e-journal] 52(6), pp.2937 - 2954. https://doi.org/10.1007/s10936-023-10031-y.

Salloum, S.A., Aljanada, R.A., Alfaisal, A.M., Al Saidat, M.R. and Alfaisal, R., 2024. Exploring the acceptance of ChatGPT for translation: An extended TAM. In: A. Al-Marzouqi, A. A. Salloum, M. Al-Saidat, A. Aburayya, and B. Gupta, eds. 2024. *Artificial Intelligence in Education: The Power and Dangers of ChatGPT in the Classroom*. Cham: Springer Nature Switzerland, pp.527 - 542. https://doi.org/10.1007/978-3-031-52280-2.

Sankaran, P., Deshbhag, R., Durbha, K., Gururajan, R. and Zhou, X., 2023. Student perceptions of ChatGPT through an expectancy value theory. In: *2023 IEEE International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. Venice, Italy, 26-29 October 2023. London: IEEE. https://doi.org/10.1109/WI-IAT59888.2023.00089.

Shang, C., Moss, A.C. and Chen, A., 2023. The expectancy-value theory: A meta-analysis of its application in physical education. *Journal of Sport and Health Science*, [e-journal] 12(1), pp.52 - 64. https://doi.org/10.1016/j.jshs.2022.01.003.

Slamet, J., 2024. Potential of ChatGPT as a digital language learning assistant: EFL teachers' and students' perceptions. *Discover Artificial Intelligence*, [e-journal] 4(1), pp.1 - 15. https://doi.org/10.1007/s44163-024-00143-2.

Solak, E., 2024. Revolutionizing language learning: How ChatGPT and AI are changing the way we learn languages. *International Journal of Technology in Education*, [e-journal] 7(2), pp.353 - 372. Available at: <https://eric.ed.gov/?id=EJ1426851> [Accessed 17 June 2024]

Sun, W., Hong, J.C., Dong, Y., Huang, Y. and Fu, Q., 2023. Self-directed learning predicts online learning engagement in higher education mediated by perceived value of knowing learning goals. *Asia-Pacific Education Researcher*, [e-journal] 32(3), pp.307 - 316. https://doi.org/10.1007/s40299-022-00653-6.

Trautwein, U., Lüdtke, O., Marsh, H.W. and Nagy, G., 2009. Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of educational psychology*, [e-journal] 101(4), pp.853 - 866. https://doi.org/10.1037/a0016306

Van Horn, K.R., 2024. ChatGPT in English language learning: Exploring perceptions and promoting autonomy in a university EFL context. *TESL-EJ*, [e-journal] 28(1), pp.1 - 26. Available at: <https://eric.ed.gov/?id=EJ1433988> [Accessed 07 June 2024]

Wigfield, A. and Eccles, J.S., 2000. Expectancy – value theory of achievement motivation. *Contemporary educational psychology*, [e-journal] 25(1), pp.68 - 81. https://doi.org/10.1006/ceps.1999.1015.

Wigfield, A., Tonks, S.M. and Klauda, S.L. (2016). Expectancy-value theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school.* 2nd ed., pp.55 - 74). Routledge.

Xiao, Y. and Zhi, Y., 2023. An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, [e-journal] 8(3), pp.1 - 12. https://doi.org/10.3390/languages8030212.

Zhu, X., Sun, H., Chen, A. and Ennis, C., 2012. Measurement invariance of expectancy-value questionnaire in physical education. *Measurement in Physical Education and Exercise Science*, [e-journal] 16(1), pp.41 - 54. Available at: <https://www.tandfonline.com/doi/abs/10.1080/1091367x.2012.639629> [Accessed 27 June 2024]

Building Resilience in Online Higher Education Facilitators: Mitigating Emotional Exhaustion

Liesl Scheepers and Geesje van den Berg
University of South Africa (UNISA), South Africa

44530684@mylife.unisa.ac.za
vdberg@unisa.ac.za

https://doi.org/10.34190/ejel.23.2.3993

Abstract: The increasing occurrence of online teaching has introduced unique challenges for facilitators, particularly concerning emotional exhaustion and burnout. The purpose of this study was to understand how online facilitators in higher education experience and navigate these challenges. The paper draws on findings presented in the author's doctoral dissertation, and explores the relationship between resilience, emotional exhaustion, and self-efficacy and how resilience can serve as a protective factor against emotional exhaustion among online facilitators at a private higher education institution in South Africa. A qualitative research approach was employed, adopting a case study design within an interpretivist paradigm. Data generation was conducted through an initial online survey sent to 1220 facilitators, from which 188 responses were received. The survey aimed to assess facilitators' levels of self-efficacy, emotional exhaustion and resilience. A purposive sampling strategy, informed by the scores achieved for each section of the survey, was used to identify ten participants for in-depth semi-structured interviews. The sample included five facilitators who demonstrated high self-efficacy and resilience, along with minimal indicators of emotional exhaustion, and five who displayed low self-efficacy levels and resilience, and significant indicators of emotional exhaustion. Thematic analysis was employed to identify key themes, focusing on work-life balance, the role of technology, adaption versus replication of teaching strategies, and the importance of resilience in maintaining engagement and effectiveness in online teaching. Findings showed that facilitators with high resilience were better able to manage emotional exhaustion. This was mainly because of their ability to establish clear work-life boundaries, use technology effectively, and adapt their teaching methods rather than replicating traditional classroom practices in an online setting. On the contrary, those with lower resilience reported greater emotional strain, challenges in adapting to online teaching, and heightened anxiety regarding the use of online tools. The findings of the study have wider implications beyond the specific institution studied, providing valuable insights for higher education institutions globally. As online education continues to grow, it is crucial to ensure that institutions provide the necessary support strategies for facilitators to build their resilience and ensure their emotional well-being in what can be a challenging teaching and learning environment.

Keywords: Resilience, Emotional exhaustion, Burnout, Self-Efficacy, Online higher education, Teaching adaptation, Work-Life balance, Faculty well-being

1. Introduction and Background

Prior to the COVID-19 outbreak, the constructs of resilience, emotional well-being, and burnout had been well-researched across a range of settings and contexts, including that of online education. With the advent of the pandemic in 2020, the focus was once again directed towards the impact of these phenomena on the lives of students and those tasked with having to support them online. While Singapore rapidly implemented measures to ensure that the needs of students were met during this time (Lim, 2020), China launched the government supported "Disrupted Classes, Undisrupted Learning" initiative (Huang, Liu, Tlili, Yang, Wang, Jemni and Burgos, 2020). The United States opted to look ahead to schools reopening and understanding what measures would need to be in place once restrictions were lifted (Minkos and Gelbar, 2021). In Australia, Green, Anderson, Tait, and Tran (2020) focused on the impact on students being isolated from campus and peer support, while Nartiningrum and Nugroho, (2020) looked at Indonesia and challenges faced regarding connectivity. Other aspects of the pandemic, such as loss of employment, isolation, or the passing of a family member, were explored by Cordaro, 2020, Ferren, 2021, and Moir, 2021. Although there were studies, such as those of Aperribai, Cortabarria, Aquirre, Verche and Borges, (2020) and Lizana, Vega-Fernandez, Gomez-Bruton, Leyton and Lera, (2021) which addressed the impact of the crisis on the family, work, and social lives of teaching staff, the attention seemed to be focused firmly on the plight of students, particularly when viewed from within the context of South Africa (Landa, Zhou, and Marongwe, 2021; Le Grange, 2021; Motala and Menon, 2020). In the wake of the pandemic, research has begun to emerge that addresses the constructs of resilience, emotional exhaustion, and burnout among educators (Malesa, 2022; Padmanabhanunni, Pretorius, Bouchard and Stiegler,

Cite this article: Scheepers, L. and van den Berg, G. 2025. "Building Resilience in Online Higher Education Facilitators: Mitigating Emotional Exhaustion", *Electronic Journal of e-Learning*, 23(2), pp 117-129, https://doi.org/10.34190/ejel.23.2.3993

2023; Spies, 2022), however, these studies have focused almost exclusively on teachers based in primary and secondary school environments rather than those facilitating in higher education settings.

2. Problem Statement

As noted, the constructs of self-efficacy and that of burnout are certainly not new. A review of the literature has revealed multiple studies on both self-efficacy and burnout among teachers, however, many of these were conducted in a pre-pandemic context. Since the advent of COVID-19, international studies have begun to emerge that investigate self-efficacy as a moderator of facilitator stress, the ability to cope with the challenges of shifting teaching practice online, and as a potential means of combatting burnout and emotional fatigue. What remains elusive, however, were any South African-based endeavours that focus on HE settings, and the potential for developing greater resilience among facilitators as a means of nurturing improved self-efficacy and emotional well-being. A study such as this is has value in that it not only adds to the narrative surrounding these constructs in HE settings, but also provides institutions with suggestions for improving resilience and self-efficacy amoung their online facilitators - whether during the next crisis, or simply in the ‘normal’ day-to-day course of events.

3. Aims and Objectives

This paper draws on findings originally presented in the author’s doctoral dissertation submitted to the University of South Africa (UNISA) in 2024, and aims to address this apparent gap by engaging with individuals in HE settings who, outside of the context of the pandemic, continue to facilitate their modules in a fully online mode of delivery. Through a better understanding of the challenges they face, institutions will be better positioned to implement strategies aimed at building resilience in these individuals, thus providing them with the necessary fortitude to stave off the onset of emotional exhaustion and burnout. This aim is in keeping with the ideas of Yildirim, Arslan and Wong (2022) who identified resilience as a critical factor in protecting against the onset of burnout, particularly in the domain of emotional exhaustion.

4. Literature Review

This literature review focuses on the notion of resilience and the construct of emotional exhaustion as experienced by online facilitators.

4.1 Defining Resilience

Definitions of resilience abound, however there are certain characteristics that are inherent in many of these. Firstly, resilience is associated with an individual’s ability to deal with threats or adversity which could potentially undermine their normal progress and development (Bertsia & Poulou, 2022, Gu & Day, 2007). Secondly, there is the suggestion that while many may have an innate measure of resilience, resilience is very much context-driven (Bertsia & Poulou, 2022, Luthar, Cicchetti & Becker, 2000) and can either be enhanced or impeded by the setting in which someone finds themselves, the people with whom they associate, and their mental fortitude. Finally, rather than being a static trait, resilience is something that can be “built up over time” (Sneha & Maheswari, 2021:33).

4.2 The Notion of Resilience

The disciplines of psychiatry and psychology were the first to grapple with the notion of resilience in relation to the ability of some children to grow and prosper despite their risk of being negatively impacted by challenging life circumstances (Howard, Dryden and Johnson, 1999). The 1980’s saw a more positive slant in the research on resilience, with the focus shifting to the strengths associated with an individual’s ability to adapt during times of crisis (Henderson and Milstein, 2003). In the years since, studies have begun to explore the “underlying protective processes” associated with resilience, in other words, how certain personal traits and protective factors might contribute to a positive outcome (Luther, Cicchetti and Becker, 2000:3). Gu and Day (2007) draw our attention to the work of Fredrickson (2004) who advocated that personal resources can be accrued, and that ‘‘through experiences of positive emotions ... people transform themselves, becoming more creative, knowledgeable, resilient, socially integrated, and healthy individuals’’ (Fredrickson, 2004:1369). Lantieri, Kyse, Harnett, Malkmus, Reevy, and Frydenberg (2011:267) add to this by suggesting that people can be taught to become more resilient through emersion in positive and supportive environments, and by those who are more “stress hardy” than themselves. As such, resilience is a dynamic construct that can be developed in individuals, increasing their ability to persevere during challenging times (Gu and Day, 2007; Howard et al., 1999; Luther et al., 2000).

4.3 Resilience Within the Context of Online Education

Duckworth (2016:2) refers to the notion of "grit", and the "determination and direction" that individuals with grit display. Naidu (2021:3) expands on this by suggesting that resilience is a product of grit, made up of "passion and perseverance", and is something that can be "developed and enhanced through direct action". He goes on to encourage educational institutions to implement proactive support strategies to foster greater resilience, particularly around the "acts of teaching and learning" (*ibid*). This suggestion is supported by the study of Liu, Zhao and Su (2022), who found that teachers with reportedly high levels of resilience were also those whose students fared better academically, and the study of Abdolrezapour, Ganjeh and Ghanbari (2023) who identified resilience as a predictor of student motivation and success in online learning contexts.

4.4 Emotional Exhaustion

Emotional exhaustion, recognised as one of the three associated domains of burnout, refers to feelings of depletion and an overextension of one's emotional reserves, as well as an increase in apathy and a reduced sense of care and concern for others (Koenig, 2014, Maslach and Leiter, 2008, Maslach, Schaufeli and Leiter, 2001). Even in a pre-pandemic context, emotional exhaustion and burnout were recognised as areas of concern for facilitators involved in online engagement. While Hislop and Ellis, (2004), noted that online engagement required more planning than contact delivery, Hogan, McKnight, and Legier (2006) highlighted the complexities associated with facilitating in an online learning environment, and Dunlap (2005) noted that the assumption of a constant online support presence risked burnout among facilitators involved in this mode of delivery. During the COVID-19 pandemic, emotional exhaustion presented as a real concern among online facilitators at a private higher education institution in South Africa, where a study by Scheepers (2024) found that even highly self-efficacious participants presented with unexpectedly elevated scores in this domain. A post-pandemic study conducted by Lucas and Vicente, (2023) which interrogated the perceived challenges experienced by 636 online facilitators across 54 countries, highlighted one of the challenges of online facilitation as being 'self-management', which includes the ability to regulate one's emotions and manage time successfully. These findings are further supported by several international studies that found that the stressors associated with moving all engagement fully online, as well as trying to cope with the demands of ensuring student wellbeing during the crisis, added to the experienced emotional exhaustion and burnout amoung facilitators (Cordaro, 2020, Kendrick, 2021, Trinidad, 2021, Yang, 2021).

4.5 Resilience and Emotional Exhaustion

De los Reyes, Blannin, Cohrssen, and Mahat (2022:50) define facilitator resilience as "the dynamic process and interaction between an academic and their everchanging environment that uses available internal and external resources to produce positive outcomes in response to different contextual, environmental, and developmental challenges". Several studies speak to the importance of building resilience amoung facilitators as a means of increasing self-efficacy and avoiding the onset of emotional exhaustion and burnout within educational settings (Beltman, 2021; Mullen, Shields and Tienken, 2021; Wang, Tsai, Lee and Ko, 2021). While Richards, Levesque-Bristol, Templin, and Graber (2016:530) suggest that "teachers who develop higher levels of resilience feel less emotionally drained, derive a greater sense of satisfaction from their work, and can interact positively with others". Resilience is not confined to defence mechanisms triggered by adversity. Recent research has shown that there are positive conditions that can shape resilience, and which have the effect of increasing the resilience of employees in organisations.

In summary, Recognising the need to leverage the lessons learnt from the pandemic, this study aimed to interrogate the notion of building resilience as a means of encouraging greater self-efficacy among online facilitators, thus providing them with an internal support mechanism that would fortify them against the demands of online engagement, whether during the next crisis or simply as a means of coping with the day-to-day demands of facilitating online.

5. Conceptual Framework

Although resilience is empirically closely related to self-efficacy, it is important to note that they are theoretically distinct (Schueler, Fritz, Dorfschmidt, Van Harmelen, Stroemer and Wessa, 2021). While resilience is closely aligned with one's ability to 'bounce back' from adversity, self-efficacy can still be present without stressors (Bertsia and Poulou, 2023). This notwithstanding, resilience and self-efficacy are often discussed in tandem, with an alignment being made between how self-efficacious someone is and their levels of persistence and resilience, suggesting that the more one is resilient and able to triumph over adversity, the more one's sense of self-efficacy is likely to be affirmed (Schwarzer and Warner, 2013). Studies have also found a significant and negative relationship between resilience and burnout (Hong, Tae and Noh, 2012; Karimi and Adam, 2018; Richards, et al., 2016), suggesting that resilient people are also far less likely to succumb to the impact of burnout. Figure 1 below illustrates the relationship between these three constructs and provides the conceptual framework for this study.

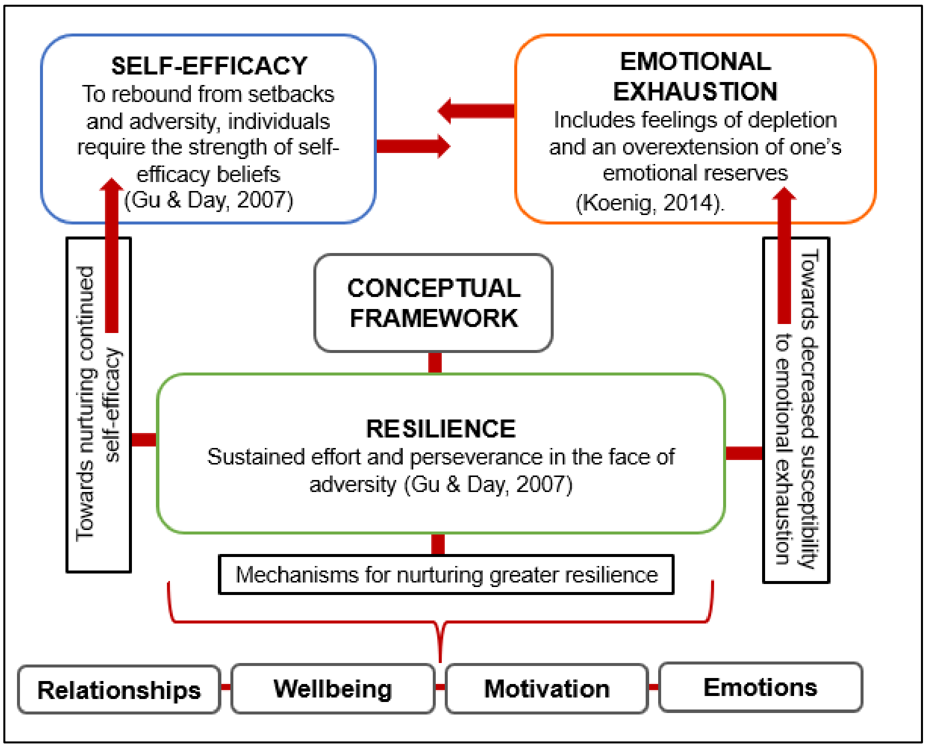

Figure 1: Conceptual Framework (Source: researchers' own)

The conceptual framework for this study highlights the interplay between these three constructs and the role that resilience might play in nurturing both improved self-efficacy, and reducing the risk of emotional exhaustion.

6. Methodology: Design and Methods

A qualitative approach was adopted for this study as it is one that allows the researcher to explore and understand the "meaning individuals or groups ascribe to a social or human problem" (Creswell, 2014:32). This qualitative study was prompted by an interest in exploring the challenges faced by online facilitators in higher education, and the role that their perceived level of resilience might play in their ability to mitigate the onset of one of the defining factors of burnout: emotional exhaustion.

A South African based higher education institution provided the setting for this study, with its facilitators comprising the population and sample. This constitutes what Creswell and Poth, (2016:61) would refer to as a "bounded system", making this a case study by design. Adopting this approach allowed the researchers to engage with participants as they shared their stories and lived experiences of what it means to engage with their students in an online environment, including the day-to-day challenges commensurate with same. For this qualitative study, an interpretivist paradigm was adopted, based on the notion that reality is socially constructed and "emerges from the way in which individuals and groups interact and experience the world" (Van Wynsberghe & Khan, 2007:90). Through the shared experiences of the participants a picture emerges of how important resilience is in mitigating the onset of emotional exhaustion, and suggests practical ways in which this can be achieved. These are more fully explored later in this study.

6.1 Population and Sample

An online survey was sent to 1 220 facilitators contracted to the selected institution, providing the population for this study. The purpose of the survey was to gather information regarding how facilitators were experiencing the demands of engaging in a fully online environment. Questions were posed that focused on their ability to adapt their teaching practice for the online environment, their confidence with leveraging digital tools for the purpose of instruction, and their perceived success in supporting students and gauging their academic success. From the survey links shared, 188 responses were received. Rather than using the data from this survey for comparative statistical analysis, the scores were used to identify potential participants for a series of semi-structured interviews. In keeping with qualitative research principles, which emphasises in-depth analysis, we felt that ten purposively selected participants was sufficient to obtain the insights and data saturation we were seeking. Prior to these engagement taking place, ethical clearance was secured from the institution (Ref:00084), and informed consent was sought and received from each of the ten participants. Participants were assured of confidentiality, and anonymity was achieved through the use of pseudonyms. As per policy, the data from the survey and semi-structured interviews has been stored on a secure institutional system.

6.2 Tools and Data Collection

The digital survey presented respondents with a series of questions under the headings of 'Self-Efficacy' and 'Emotional Exhaustion'. Responses were scored between 0 to 4 in both sections, with zero being attributed to '*I could do nothing'*, and four being "*I could do a great deal*" in the section focusing on self-efficacy, and zero for "*I felt highly motivated*", to four for "*I felt overwhelmed*" in the section focusing on emotional exhaustion. Table 1 provides an example of some of the questions posed in the survey, possible responses, and how these were scored.

Table 1: Examples of survey questions, responses and scores (Source: researchers' own)

Self-Efficacy	*To what extent are you able to navigate the institutional LMS to successfully facilitate your module/s online?*	*To what extent are you able to navigate the internet in order to provide links and/or additional resources for your online students?*	*To what extent are you able to use your synchronous sessions to maximise interaction between students?*	**SCORE**
Example Only	I can do a great deal (4)	I can do quite a bit (3)	I can do some (2)	9
Example Only	I can do nothing (0)	I can do very little (1)	I can do some (2)	3
Emotional Exhaustion	I sometimes have trouble sleeping and wake up feeling exhausted	I find it hard to concentrate and can be absentminded	Between the demands of teaching online and my family – I have very little left to give	**SCORE**
Example Only	I never feel like this (0)	I feel like this a few times (1)	I feel like this once a week (2)	3
Example Only	I feel like this once a week (2)	I feel like this a few times each week (3)	I feel like this every day (4)	9

Based on the scores achieved, the sample group comprised five individuals who had presented with low levels of resilience and self-efficacy, and significant indicators of emotional exhaustion, and five individuals with high self-efficacy scores, and minimal indicators of emotional exhaustion. The data gleaned from their survey responses was used to guide the engagement during the interviews, with participants being asked to comment on the scores achieved or to elaborate on certain questions in more detail. For this particular study, the age, ethnicity, tenure, and online teaching experience of the participants were not considered, instead the researchers were interested in the scores achieved for the survey, and the participants' willingness to share their experiences of facilitating their modules in a fully online context.

Prior to the semi-structured interviews, the questions were piloted with a group of volunteers to determine clarity of wording and intent. Based on the feedback received, question nine was amended to ensure the validation of experienced emotions rather than diminishing the individual's experience in any way, however unintentional. This amendment is presented in table 2 below.

Table 2: Amendment to question nine after pilot exercise (Source: researchers' own).

Interview question 9	
Original	"*Did you find that you were ever irrational and quick to anger*",
Amended to	"*Did you ever experience heightened emotions that threatened to get the better of you?*".

Through detailed inductive thematic analysis of the participants' shared experiences, themes were sought that would guide the reader through a seamless flow of inductive reasoning (Fouché, 2021, Malakar, 2022).

Transcripts of the interviews were manually transcribed to allow for a deeper emersion in each of the stories, and to ensure that the voices of the participants directed the course of the narrative (Maree et al., 2016, Terrel, 2016). These transcripts were then re-read over several iterations to define and refine the final three sub-themes which served to address the main theme of determining what can be done to minimise emotional exhaustion and promote resilience among online facilitators in higher education.

6.3 Questions Asked and Themes Explored

As noted, the overarching intention behind this study was to determine what might cause emotional exhaustion in online facilitators within a higher education setting and how improving their levels of resilience might prevent them from experiencing this phenomenon in the future. Five purposively selected participants who had presented with high levels of self-efficacy and resilience were encouraged to share their experiences of facilitating online and why they believe they have been able to avoid succumbing to the symptoms associated with burnout, particularly that of emotional exhaustion. In tandem with this, five other participants who had presented with low levels of resilience and self-efficacy were asked to share their stories, and what impact their experience of emotional exhaustion has had on their lives and their engagement with their students.

6.4 Context and General Participant Information

The institution at which this study was conducted was first established in 1991. In the 33 years since its inception, it has grown to comprise nine contact campuses across the country, and an online centre established in 2017. It is the experiences of the facilitators tasked with supporting these online students that was of particular interest to the researchers.

As noted, a survey was distributed with the intention of purposively identifying ten individuals who would then participate in one-on-one semi-structured interviews. The tenure of the ten participants in this study ranged from five to 19 years in higher education, with each having had at least two years' experience in online facilitation. The participants comprised seven women and three males, aged 36 to 'older than 55'. Areas of specialisation in terms of subject matter expertise included the Social Sciences, Finance and Accounting, Business Management, and Communication Science. It is relevant to note that although specific modules facilitated by each participant were not intentionally highlighted during the interviews, some reference has been made to these in order to provide a context for certain participant quotes included in the discussion. Pseudonyms were used to protect their identity. A summary of these participants is provided in Table 3 below.

Table 3: A summary of study participants and their scores (Source: researcher's own)

Group One	Participants				
Low Self-Efficacy / High Emotional Exhaustion	Siva	Hannah	Alison	Malcolm	Nelly
Self-Efficacy / 50	8	12	18	18	15
Emotional Exhaustion / 20	14	14	12	14	14
Group Two	Participants				
High Self-Efficacy / Low Emotional Exhaustion	Layla	Craig	Paula	Tamika	Louise
Self-Efficacy / 50	37	31	36	31	36
Emotional Exhaustion / 20	6	6	6	9	8

7. Data Analysis And Discussion

Through a process of reiterative reading and thematic analysis, four sub-themes were identified to support the main theme of determining how to minimise emotional exhaustion and promote resilience among online facilitators in higher education.

7.1 Sub-themes

1. Balance through boundaries
2. Technology as help or hinderance
3. Adaptation rather than replication
4. Resilience is the key

Each of these are discussed from the perspective of those more self-efficacious and resilient participants who appear to thrive in an online teaching and learning environment, and those with lower levels of self-efficacy and reduced resilience who often find the challenges of facilitating in this environment to be emotionally exhausting.

7.2 Sub-Theme 1: Balance Through Boundaries

A study by Bauwens, Muylaert, Clarysse, Audenaert, and Decramer, (2020:3) suggests that "depending on how employees manage the boundaries between their work and life domains, activities in one domain can create spillovers to the other domain, resulting in role conflict or role confusion". Similarly, participants in this study who shared that they experienced high levels of emotional exhaustion were also those who admitted that they often allowed the boundaries to blur between their responsibilities outside of the online teaching environment and their commitment to supporting their students. When participants in this study were asked about their ability to maintain a work-life balance while supporting their students through online facilitation, each of the ten highlighted the importance of achieving their own sense of balance on a daily basis; with some sharing the negative impact that an absence of this balance can have on their lives. One participant (Nelly) shared that if she does not plan her life, and put boundaries in place, she "is a mess", with hours and days blurring, and a compounded sense of not "getting to everything that needs taking care of". Hannah, another participant, explained that she found maintaining a day-to-day balance to be a challenge, and that often the demands of work, family, and other external commitments "collide".

Khateeb (2021:28), citing Kirchmeyer (2000), suggests that work-life balance is the "achievement of fulfilling experiences in the different aspects of life that require various resources, like energy, time and commitment, and these resources are spread across all the domains". This balance is not always easy to achieve, especially for those facilitating in higher education settings where the prevalence of technologies creates an environment where facilitators are potentially accessible to their students at any time of the day or night (Ilić-Kosanović, 2021; Parham & Rauf, 2020). As two participants shared:

> ***Alison**: It can all just become too much, you know. When a student messages me I feel they need me to respond then and there. I know I would want a quick response so I feel like I cannot make them wait. It can be quite exhausting to keep up.*
>
> ***Siva**: I can't switch off. I will check my WhatsApp after hours, and if there is a question, I feel like I need to answer it right away. It can sometimes go on quite late, especially when assessments are due.*

The Bauwens et al., (2020) study goes on to explain that individuals who capitalise on the flexibility afforded by technology are also those who are able to reduce the demands made on them when facilitating in an online environment and ensure a greater sense of equilibrium in their work-life balance. The central role that technology inevitably plays in an online learning environment was supported by the data in this study and gave rise to the second sub-theme.

7.3 Sub-Theme 2: Technology as Help or Hindrance

Shifting one's facilitation online requires significant changes to one's teaching practice, often demanding "extra planning and effort" (Lucas & Vicente, 2023:5092). Lee and Ogawa (2021) suggest that where an institution has a track record of having promoted the use of teaching technologies, facilitators may develop a level of familiarity that will allow them to adapt to an online learning environment with relative ease. The institution at which this study was conducted implemented a Learning Management System (LMS) in 2014. In the years since, training has been ongoing, focusing on developing blended and online teaching practices. Despite this, the findings in this study still suggest a significant divide between those more efficacious participants who have enthusiastically embraced the use of technology for their online engagement and others who, despite prior exposure to these same technologies, find this requirement intimidating and stressful resulting in what Ventura, Salanova, and Llorens, (2015:280) refer to as a "crisis in efficacy".

> ***Nelly**: I don't mind the discussion board tool, but … setting up 'auto-graded tasks', there I just get scared that I am going to push the wrong button or create something that doesn't work properly when my students try it, so I just stick to what I know I can do.*
>
> ***Alison**: I'm not sure how to use [the tools] properly, to get the most out of them. Obviously, I know the basics; I know how to set up and run a Collab session, I know how to put up additional resources, but I'm not sure how to really leverage the tools as I would like… and I find this incredibly stressful.*

While Nelly and Alison share their reticence concerning technology, Paula and Layla appear to have adopted a more intentional approach to adopt the available technologies commensurate with the requirements of online engagement:

> ***Paula****: I love technology, it's always been a passion… I suppose I see technology as a challenge rather than a hardship.*

> ***Layla:*** *I like to figure out how a tool works and whether it would serve a purpose for me in my module. Same with the internet, I don't let it phase me, I also refused to let it beat me (laughs).*

Whether considered a help or a hindrance, the reality remains that facilitating online requires the use of technologies to support the student learning journey. What became evident during an analysis of the data was that the greatest anxiety and emotional exhaustion was experienced by those participants who had attempted to use the technology to simply replicate their contact classroom practices rather than interrogating how these might need to be adapted for an online setting. This discovery led to the third sub-theme.

7.4 Sub-Theme 3. Adaptation Rather Than Replication

A literature review conducted by Fernández-Batanero, Román-Graván, Reyes-Rebollo, and Montenegro-Rueda (2021) highlighted that facilitators who felt that they had not developed the necessary online pedagogies were likely to feel threatened by the perceived challenges associated with shifting their teaching practice online (Lee and Tsai, 2010). This lack of familiarity can lead to increased anxiety, emotional exhaustion, and even complete avoidance regarding the use of online teaching technologies (McIlroy and Bunting, 2002). A study by Ma, Chutiyami, Zhang, and Nicoll (2021), noted similar findings and explained that when facilitators attempted to use technology to simply replicate their contact teaching practices in the online environment, their levels of anxiety and burnout increased. Similarly, in this study, it was those participants who appeared reluctant, or unable, to make the necessary changes to their practice that experienced heightened levels of emotional exhaustion and a sense of disengagement from their students.

> ***Malcolm:*** *I didn't really understand how to adapt, or what I needed to adapt, or what 'adapting' actually meant. [So, I] just did things the same way as we did in the classroom, but that obviously wouldn't work… I can see that now, but not [when I started].*

> ***Hannah****: If you want to write out actual formulas and calculations of numbers with a lot of lines, you can't read it on that digital whiteboard. It really doesn't translate well, so … I tend to avoid it.*

> ***Siva****: Online [students] have no idea who you are, and the truth is they need you to be different online, and I did not see that in the beginning.*

Where participants had understood from the outset that contact teaching practices do not, and, one could perhaps argue, should not directly translate into an online setting, a conscious decision had been made to adapt their approach to one more suited to fully online engagement.

> ***Louise****: I use a mix of audio, visual and written resources and carefully plan these so that it [isn't] too…overwhelming… but rather a complementary combination.*

> ***Tamika****: In online [facilitation], it needs to be far more 'flipped'; students need to get ahead with the content and then use those online sessions to question, consolidate or unpack what they have covered for themselves.*

> ***Paula****: I make a clear distinction between synchronous and asynchronous sessions. I create short pre-recordings that [students] can watch in their own time, and then the [live] sessions are for sharing, talking…explaining concepts and so on.*

During the interviews, the notion of resilience was raised by seven of the ten participants, thus resulting in the final sub-theme.

7.5 Sub-theme 4. Resilience is the key

While participants like Louise, Paula, and Tamika associated their ability to meet the demands of online facilitation with their sense of resilience and a willingness to adapt, others like Nelly and Hannah commented on how they felt that their lack of resilience was a contributing factor to them sometimes feeling overwhelmed and emotionally exhausted by the requirements of the online environment.

Paula: *I am quite tough, or should I say resilient, when it comes to challenges that are thrown my way. Changing to online was a big adjustment, but I was excited to take it on.*

Louise*: Oh, I am definitely resilient. Very little phases me. My life has been about change, so I knew I could take on teaching in a new environment [like online].*

Nelly*: I can be a bit of a mess when I take on too much at once, and moving my teaching online was a massive step that honestly left me quite shattered most days.*

Hannah: *I wish I was tougher, more resilient, but the truth is I'm just not. I can become overwhelmed quite quickly, especially when the challenge seems a bit larger than life.*

What each of these participants have shared is in keeping with the work of Bonanno (2021:5) who suggest that resilience "is not a simple, one-dimensional construct but, rather, involves multiple interacting components". To this he adds that certain personality traits or dimensions have been empirically linked to resilience, albeit with modest effects. Traits such as optimism, coping self-efficacy, and being challenge-orientated, can result in the kind of flexible mindset that allows certain individuals to cope with challenges better than others are able to do (*ibid*). Gu and Day (2013), concur, positing that resilience is a complex process that considers a range of personal and contextual factors. It is with this in mind that the following recommendations are made. This means that when institutions look to ways of fostering increased resilience among their facilitators, it is important to understand the process and to recognise that resilience can be used to encourage agency and action, and strengthen resolve (Keogh, Garvis, Pendergast and Diamond, 2012; Nandy, Lodh and Tang, 2021).

8. Recommendations for Institutions

Based on the findings of this study, the following three recommendations are made:

- **1. Academic support**

Institutions are encouraged to proactively establishment of online communities of practice at the beginning of the academic year within the institution's LMS, creating virtual spaces where new and existing facilitators can be directed to connect with peers and to share ideas and resources.

- **2. Peer Support**

Peer support that acts as a buddy-system where support is available for the day-to-day challenges that can present in an online learning context, the sharing of suitable resources, and supporting those individuals less confident in the use of technology to facilitate their practice and encourage a better understanding of what it means to engage with content online.

- **3. Institutional Leadership**

Leadership are encouraged to create at least two 'open mic' type opportunities per semester which allow for more informal engagement between leadership and those tasked with facilitating the institution's qualifications. Establishing this type of forum would serve to minimise the perceived distance between leadership and facilitators, promote greater agency, and allow leadership to gain a richer understanding of the team of facilitators who represent them online and in the classroom.

These recommendations are in keeping with the ideas of Gu and Day (2007) who suggest that by focusing on building and maintaining teacher resilience there will be a direct and positive impact on the quality of teaching as well as the academic achievements of students, while Mullen et al., (2021:14) highlight that an institution that sets high standards, provides clear administrative goals, and allows for "meaningful participation of teachers in decision-making" will create an environment in which teachers can experience a greater sense of agency and purpose leading to increased resilience, improved morale, and reduced susceptibility to burnout.

9. Limitations and Recommendations for Further Study

This study was positioned within the context of private higher education in South Africa, thus offering a slightly limited, albeit unique, lens through which the constructs of resilience and emotional exhaustion among online facilitators were investigated. Further research that interrogates these same constructs as experienced by online facilitators at public higher education institutions would provide a broader perspective and offer valuable insights into how institutions might mitigate the impact of the day-to-day challenges beyond the scope of a pandemic faced by online facilitators across a range of settings.

While information regarding the age, ethnicity, tenure, and online teaching experience of the participants was gathered using an online survey, for the purpose of this study, this data was not integral to the analysis that was undertaken. Further studies that interrogate whether an online facilitator's level of self-efficacy and emotional exhaustion have any correlation to their age, ethnicity, tenure, and/or teaching experience would add another dimension to understanding these constructs and how institutions might go about developing resilience-focused support strategies that are more closely aligned to the needs of the individual.

Current research suggests that it is not unrealistic to anticipate some form of crisis in the future that would require educational institutions to apply strategies not dissimilar to those implemented in 2020 and 2021 (Desmond-Hellmann, 2020, Hixon, 2020, Nandy, et al., 2021). What is important is that these institutions are better prepared than they have previously been, suggesting a measure of urgency for understanding how to build greater resilience among teaching staff in order to mitigate the onset of emotional exhaustion that can be brought about by stressful situations that might arise, even outside the context of a pandemic (Ross, Scanes and Locke, 2023).

10. Conclusion

A review of the literature cited in this study suggests that while much has been written about the constructs of resilience and emotional exhaustion among teachers in schools, more research is required that interrogates these same phenomena among online facilitators in higher education settings. By engaging with online facilitators working within this context, and hearing their stories regarding the day-to-day challenges associated with this mode of instruction, this study aimed, at least in part, to answer this call. Through these shared insights, much can be learnt about how one's sense of resilience, or perhaps lack thereof, might impact whether burnout is experienced, with particular reference to the domain of emotional exhaustion. The findings gained from these interactions can guide institutions in their plans to proactively support their facilitators in becoming more resilient to the numerous factors at play within a higher education online setting. Resilience, the researchers would argue, should not be a resource limited to a select few, but rather one that every online facilitator has the opportunity to develop.

AI Declaration: The use of AI has not formed any part of this research and/or the drafting of this article.

Ethics Statement: All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of The Independent Institute of Education (R.00084 [REC]).

References

Abdolrezapour, P., Jahanbakhsh Ganjeh, S. and Ghanbari, N., 2023. Self-efficacy and resilience as predictors of students' academic motivation in online education. *Plos one*, *18*(5), p.e0285984. https://doi.org/10.1371/journal.pone.0285984.

Aperribai, L., Cortabarria, L., Aguirre, T., Verche, E. and Borges, Á., 2020. Teacher's physical activity and mental health during lockdown due to the COVID-2019 pandemic. *Frontiers in Psychology,* 11, p.577886. http://doi.org/10.3389/fpsyg.2020.577886.

Bauwens, R., Muylaert, J., Clarysse, E., Audenaert, M. and Decramer, A., 2020. Teachers' acceptance and use of digital learning environments after hours: Implications for work-life balance and the role of integration preference. *Computers in Human Behavior, 112,* p.106479. https://doi.org/10.1016/j.chb.2020.106479.

Beltman, S., 2021. Understanding and examining teacher resilience from multiple perspectives. *Cultivating teacher resilience*, pp.11-26. https://doi.org/10.1007/978-981-15-5963-1.

Bertsia, V. and Poulou, M., 2023. Resilience: Theoretical Framework and Implications for School. *International Education Studies*, *16*(2), p.1-12. https://doi.org/10.5539/ies.v16n2p1.

Bonanno, G.A., 2021. The resilience paradox. *European journal of Psychotraumatology*, *12*(1), p.1942642. https://doi.org/10.1080/20008198.2021.1942642.

Cordaro, M., 2020. Pouring from an empty cup: The case for compassion fatigue in higher education. *Building Healthy Academic Communities Journal, 4*(20), p.17-28. https://doi.org/10.18061/bhac.v4i2.7618.

Creswell, J.W., 2014. Research Design: Qualitative, Quantitative, and Mixed Methods Approached. 4th Edition. SAGE Publications, 2008. Educational Research. New Jersey: Pearson Education

Creswell, J.W. and Poth, C.N., 2016. *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

De los Reyes, E.J., Blannin, J., Cohrssen, C. and Mahat, M., 2022. Resilience of higher education academics in the time of the 21st century pandemics: a narrative review. *Journal of Higher Education Policy and Management, 44*(1), p. 39-56. https://doi.org/10.1080/1360080X.2021.1989736.

Desmond-Hellmann, S. 2020. Preparing for the next pandemic. *The Wallstreet Journal,* <Available at: https://www.wsj.com/articles/preparing-for-the-next-pandemic-11585936915. [Accessed 19 January 2025].

Dimakos, I. and Papakonstantinopoulou, A., 2012. Providing psychological and counselling services to Roma students: A preliminary report for a three-year longitudinal project. *Creating communities: Local, national and global*, pp.94-103. ISBN 978-1-907675-19-5.

Duckworth, A., 2016. *Grit: The power of passion and perseverance* (Vol. 234). New York, NY: Scribner. https://doi.org/10.1111/peps.12198.

Dunlap, J.C., 2005. Problem-based learning and self-efficacy: How a capstone course prepares students for a profession. *Educational Technology Research and Development*, *53*(1), pp.65-83. ISSN 1042–1629.

Ferren, M., 2021. Social and emotional Support for Educators During and After the Pandemic. *Centre for American Progress: ERIC.*

Fernandez-Batanero, J.M., Roman-Gravan, P., Reyes-Rebollo, M.M. and Montenegro-Rueda, M., 2021. Impact of educational technology on teacher stress and anxiety: A literature review. *International journal of environmental research and public health*, *18*(2), p.548. https://doi.org/10.3390/ijerph18020548.

Fouché, C. B. (2021). Introduction to the research process. In Fouché, C.B., Strydom, H. & Roesterburg, W. J. H. Research at grassroots for the social sciences and human services professions (5th Edition). Pretoria: Van Schaik.

Fredrickson, B.L., 2004. The broaden–and–build theory of positive emotions. *Philosophical transactions of the royal society of London. Series B: Biological Sciences*, *359*(1449), 1367-1377. https://doi.org/10.1098/rstb.2004.1512.

Green, W., Anderson, V., Tait, K., and Tran, L.T., 2020. Precarity, fear and hope: reflecting and imagining in higher education during a global pandemic. *Higher Education Research & Development, 39*(7), 1309-1312. https://doi.org/10.1080/07294360.2020.1826029.

Gu, Q. and Day, C., 2007. Teachers' resilience: A necessary condition for effectiveness. *Teaching and Teacher education*, *23*(8), pp.1302-1316. https://doi.org/10.1016/j.tate.2006.06.006.

Gu, Q. and Day, C., 2013. Challenges to teacher resilience: Conditions count. *British educational research journal*, *39*(1), pp.22-44. https://doi.org/10.1080/01411926.2011.623152.

Henderson, N. and Milstein, M.M., 2003. *Resiliency in schools: Making it happen for students and educators*. Corwin press.

Hislop, G.W. and Ellis, H.J., 2004. A study of faculty effort in online teaching. *The Internet and higher education*, *7*(1), pp.15-31. https://doi.org/10.1016/j.iheduc.2003.10.001.

Hixon, T. 2020. Get Ready To Live With COVID-19. *Forbes,* [online] Available at: https://www.forbes.com/sites/toddhixon/2020/03/12/get-readyto-live-with-covid-19/#5bd55a094782 [Accessed 24 October 2024].

Hogan, R.L. and McKnight, M.A., 2007. Exploring burnout amoung university online instructors: An initial investigation. *The internet and higher education, 10*(2), pp.117-124. https://doi.org/10.1016/j.iheduc.2007.03.001.

Hong, M.J., Tae, Y.S. and Noh, M.Y., 2012. Relationships between stress, ways of coping and burnout of family caregivers of cancer patients. *Asian Oncology Nursing*, *12*(1), pp.92-99. https://doi.org/10.5388/aon.2012.12.1.92.

Howard, S., Dryden, J. and Johnson, B., 1999. Childhood resilience: Review and critique of literature. *Oxford review of education*, 25(3):307-323. https://doi.org/10.1080/030549899104008.

Huang, R.H., Liu, D.J., Tlili, A., Yang, J.F., Wang, H.H., Jemni, M., Burgos, D. 2020. Handbook on Facilitating Flexible Learning During Educational Disruption: The Chinese Experience in Maintaining Undisrupted Learning in COVID-19 Outbreak. Beijing: Smart Learning Institute of Beijing Normal University. http://creativecommons.org/licenses/by-sa/3.0/igo/.

Ilić, K.T., 2021. Work-life balance of teaching staff at higher education during COVID 19 pandemic. *Serbian Journal of Engineering Management*, *6*(1), 63-71. https://doi.org/ 10.5937/SJEM2101064I.

Karimi, M.N. and Adam, S.B., 2018. A structural equation modelling analysis of the relationships between perceived occupational stress, burnout, and teacher resilience. *Journal of Second Language Teacher Education, 1*(1), 49-72. https://doi.org/10. 1558/slte.27069.

Kendrick, A., 2022. Compassion fatigue, burnout, and the emotional labor of educational workers. *The International Journal of Health, Wellness and Society*, *13*(1), p.31. https://doi.org/10.18848/2156-8960/CGP/v13i01/31-55.

Keogh, J., Garvis, S., Pendergast, D. and Diamond, P., 2012. Self-determination: Using agency, efficacy and resilience (AER) to counter novice teachers' experiences of intensification. *Australian Journal of Teacher Education, 37*(8), 46-65. https://doi.org/10.1422/AJTE.2021.2021V37N9.3.

Khateeb, F.R., 2021. Analysis Of Work-Life Conflict Construct In Case Of Telework-A Case Study Of Teachers From Kashmir Valley (India) During Covid-19 Pandemic. *SEA–Practical Application of Science*, *9*(26), pp.143-154.

Koenig, A., 2014. Learning to prevent burning and fatigue: Teacher burnout and compassion fatigue. Electronic Thesis and Dissertation Repository, Paper 1928. https://ir.lib.uwo.ca/etd/1928

Landa, N., Zhou, S. and Marongwe, N. 2021. Education in emergencies: Lessons from COVID-19 in South Africa. *International Review of Education, 67*(1-2), pp.167-183. https://doi.org/10.1007/s11159-021-09903-z.

Lantieri, L., Kyse, E.N., Harnett, S., Malkmus, C., Reevy, G.M. and Frydenberg, E., 2011. Building inner resilience in teachers and students. *Personality, stress, and coping: Implications for education*, pp.267-292.

Lee, S.C.N. and Ogawa, C., 2021. Online teaching self-efficacy—how English teachers feel during the Covid-19 pandemic. *Indonesian TESOL Journal*, *3*(1), pp. 1-17.

Lee, M.H. and Tsai, C.C. 2010. Exploring teachers perceived self-efficacy and technological pedagogical content knowledge with respect to educational use of the world wide web. *Instructional Science, 38*(1), pp.1-21. https://doi.org/10.1007/s11251-008-9075-4.

Le Grange, L. 2021. Covid-19 pandemic and the prospects of education in South Africa. *Prospects,* 51, pp.425-436. https://doi.org/10.1007/s11125-020-09514-w.

Lim, M., 2020. Educating despite the Covid-19 outbreak: lessons from Singapore. *Times Higher Education,* 20. THE World Universities Insights Limited, London, UK.

Liu, Y., Zhao, L. and Su, Y.S., 2022. The impact of teacher competence in online teaching on perceived online learning outcomes during the COVID-19 outbreak: A moderated-mediation model of teacher resilience and age. *International journal of environmental research and public health*, *19*(10), p.6282. https://doi.org/10.3390/ijerph19106282.

Lizana, P.A., Vega-Fernandez, G., Gomez-Bruton, A., Leyton, B. and Lera, L., 2021. Impact of the COVID-19 Pandemic on Teacher Quality of Life: A Longitudinal Study from before the Health Crisis. *International jounral of environmental research and public health, 187*(7), 3764. https://www.mdpi.com/1660-4601/18/7/3764#.

Lucas, M. and Vicente, P.N., 2023. A double-edged sword: Teacher's perceptions of the benefits and challenges of online teaching and learning in higher education. *Education and Information Technologies, 28*(5), pp.5083-5103..

Luthar, S.S., Cicchetti, D. and Becker, B., 2000. Research on resilience: Response to commentaries. *Child development*, *71*(3), 573-575. https://doi.org/10.1111/1467-8624.00168.

Ma, K., Chutiyami, M., Zhang, Y., Nicoll, S., 2021. Online Teaching Self-Efficacy during COVID-19: Changes, Its Associated Factors and Moderators. *Education Information Technology,* pp. 226:1-23. https://doi.org/10.1007/s10639-021-10486-3.

Malakar, I. M., 2022. Carving the research proposal. *Journal of population and Development, 3*(1), pp.42-63. https://doi.org/10.3126/jpd.v3i1.48805.

Malesa, M., 2022. *The impact of emotional labour on teachers in public and private schools during the COVID-19 pandemic* (Doctoral dissertation, University of Johannesburg).

Maree, K., Creswell, J.W., Ebersohn, L., Eloff, I., Ferreira, R., Ivankova, N.V., Jansen, J.D., Nieuwenhuis, J., Pietersen, J. and Clark, V.L.P. 2016. First steps in research. K. Maree (Ed.) Pretoria: Van Schaik Publishers.

Maslach, C. and Leiter, M.P., 2008. Early predictors of job burnout and engagement. *Journal of applied psychology*, *93*(3), p.498. https://doi.org/ 10.1037/0021-9010.93.3.498.

Maslach, C., Schaufeli, W.B. and Leiter, M.P. 2001. Job burnout. *Annual Review of Psychology, 52*(1), pp.397-422. https://doi.org/10.1146/annurev.psych.52.1.397.

McIlroy, D. and Bunting, B. 2002. Personality, behaviour and academic achievement: Principles for educators to inculcate and students to model. *Contemporary Educational Psychology, 27*(2), pp.326-337. https://doi.org/10.1006/ceps.2001.1086.

Minkos, M.L. and Gelbar, N.W., 2021. Considerations for educators in supporting student learning in the midst of Covid-19. *Psychology in the Schools, 58*(2), 416-426. https://dpoi.org/10.1002/pits.22454.

Moir, J., 2021. Social Emotional Learning: A Discourse for the Times. *Proceedings of IAC 2021 in Vienna,* ISBM 978-80-88203-22-3.

Motala, S. and Menon, K. 2020. In search of the 'new normal': Reflections on teaching and learning during Covid-19 in a South African university. *Southern African Review of Education with Education Production, 26*(1), pp.80-99. https://hdl.handle.net/10520/ejc-sare-v26-n1-a6.

Mullen, C.A., Shields, L.B. and Tienken, C.H., 2021. Developing teacher resilience and resilient school cultures. *Journal of Scholarship & Practice*, *18*(1), pp.8-24.

Naidu, S., 2021. Building resilience in education systems post-COVID-19, *Distance Education*, *42* (1), pp.1-4. https://doi.org/10.1080/01587919.2021.1885092.

Nandy, M., Lodh, S. and Tang, A., 2021. Lessons from Covid-19 and a resilience models for higher education. *Industry and Higher Education, 35*(1), pp.3-9. https://doi.org/10.1177/0950422220962696.

Nartiningrum, N., and Nugroho, A., 2020. Online Learning amidst Global Pandemic: EFL Students' Challenges, Suggestions, and Needed Materials. *Academic Journal of English Language Education, 4*(2), 115-140. https://doi.org/10.29240/ef.v4i2.1494.

Padmanabhanunni, A., Pretorius, T.B., Stiegler, N. and Bouchard, J.P., 2022. A serial model of the interrelationship between perceived vulnerability to disease, fear of COVID-19, and psychological distress among teachers in South Africa. In *Annales Medico-psychologiques, revue psychiatrique. 180*(1), pp.23-28. Elsevier Masson. https://doi.org/10.1016/j.amp.2021.11.007.

Parham, S. and Rauf, M.A., 2020. COVID-19 and obligatory remote working in HEIs: An exploratory study of faculties' work-life balance, well-being and productivity during the pandemic. *International Journal of Economics, Commerce and Management*, *8*(12), pp.384-400. http://ijecm.co.uk/wp-content/uploads/2020/12/81223.pdf.

Richards, K.A.R., Levesque-Bristol, C., Templin, T.J. and Graber, K.C., 2016. The impact of resilience on role stressors and burnout in elementary and secondary teachers. *Social Psychology of Education,* (19), pp.511-536. https://doi.org/10.1007/s11218-016-9346-x.

Ross, P.M., Scanes, E. and Locke, W., 2023. Stress adaptation and resilience of academics in higher education. *Asia Pacific Education Review,* pp:1-21. https://doi.org/10.1007/s12564-023-09829-1.

Scheepers, L.H., 2024. *The connection between open distance learning facilitators' sense of self-efficacy and experienced burnout during a time of crisis.* (Doctoral thesis, University of South Africa).

Schueler, K., Fritz, J., Dorfschmidt, L., Van Harmelen, A.L., Stroemer, E. and Wessa, M., 2021. Psychological network analysis of general self-efficacy in high vs. low resilient functioning healthy adults. *Frontiers in psychiatry*, 12, p.736147. https://doi.org/10.1111/ecc.12952.

Schwarzer, R. and Warner, L.M., 2013. Perceived self-efficacy and its relationship to resilience. In S. Prince-Embury & D. H. Saklofske (Eds.), The Springer series on human exceptionality: Resilience in children, adolescents, and adults: Translating research into practice (pp. 139-150). https://doi.org/10.1007/978-1-4614-4939-3_10.

Sneha, K. and Maheswari, G.S., 2021. Short Communication Five-Pronged Approach To Building Resilience AT Workplace Teacher's Community In Andaman Island During COVID-19. Journal of Strategic Human Resource Management, 10 (3). ISSM"2277-2138.

Spies, J., 2022. *Exploring Teacher Experiences of Subjective Well-Being During a Pandemic* (Master's thesis, University of Pretoria (South Africa)).

Terrel, S.R. 2016. *Writing a Proposal for Your Dissertation*. Guilford Press (E Book).

Trinidad, J.E., 2021. Teacher satisfaction and burnout during COVID-19: what organizational factors help?. *International Journal of Leadership in Education*, pp.1-19. https://doi.org/10.1080/13603124.2021.2006795.

Van Wynsberghe, R. and Khan, S., 2007. Redefining case study. *International journal of qualitative methods, 6*(2), pp.80-94. https://doi.org/10.1177/160940690700600208.

Ventura, M. Salanova, M. and Llorens, S., 2015. Professional self-efficacy as a predictor of burnout and engagement: The role of challenges and hinderance demands. *The Journal of psychology, 149*(3), pp.277-302. https://doi.org/10.1080/00223980.2013.876380.

Wang, I.A., Tsai, H.Y., Lee, M.H. and Ko, R.C., 2021. The effect of work–family conflict on emotional exhaustion and job performance among service workers: The cross-level moderating effects of organizational reward and caring. *The International Journal of Human Resource Management*, *32*(14), pp.3112-3133. https://doi.org/10.1080/09585192.2019.1651373.

Yang, C., 2021. Online Teaching Self-Efficacy, Social-Emotional Learning (SEL) Competencies, and Compassion Fatigue Among Educators During the COVID-19 Pandemic. *School of Psychology Review,* https://doi.org/10.1080/2372966X.2021.1903815.

Yıldırım, M., Arslan, G. and Wong, P.T., 2022. Meaningful living, resilience, affective balance, and psychological health problems among Turkish young adults during coronavirus pandemic. *Current Psychology*, *41*(11), pp.7812-7823. https://doi.org/10.1007/s12144-020-01244-8.

Exploring Determinants of Online Learning Acceptance: The Role of Readiness, Peer Support, and Instructional Support

Leroy Robinson Jr.

Department of Decision Sciences, Economics, Finance and Marketing, University of Houston-Clear Lake, Houston, TX, USA

robinsonjr@uhcl.edu

https://doi.org/10.34190/ejel.23.2.4048

Abstract: This study investigates influential aspects of students' acceptance of online learning: students' readiness, peer support, and instructional support. Readiness, driven by motivation, technical skills, and self-management, affects online learning participation. This readiness is made up of self-learning and communication abilities in online environments, which influence acceptance and satisfaction. Peer support inspires collaboration and enhances learning gains. It reduces loneliness, boosts motivation, and facilitates teamwork, while instructional support aids learning through organized interaction and feedback. Instructional support (e.g., immediate feedback and well-structured instruction) also improves engagement and perceived accomplishment in online education. The study employed a systematic sampling plan, using 308 students at an American university pursuing online or blended courses. Confirmatory factor analysis was used to validate the measurement model, confirming construct validity and reliability. Structural equation modeling confirmed the tested hypotheses and relationships among the variables. Students who perceive online learning as useful and convenient are more inclined to engage with the learning management system, which aligns with the technology acceptance model. Psychological, technological, and behavioral readiness play a primary role in determining whether self-efficacious and self-regulated learners will adapt. Peer support is important in offsetting the alienating effect of e-learning and facilitating engagement, motivation, and cognitive presence. Interpersonal behavior, such as peer mentoring and group discussion, increases social belonging, reduces anxiety, and develops academic resilience. Instructional support is critical to the acceptance of online learning. Timely scaffolding and immediate feedback increase students' engagement and motivation. Institutional investments in technical and non-technical resources enable active participation. The study's broader implications are multifaceted and require a holistic approach focusing on content delivery, actively preparing students, fostering social connections, and supporting them throughout their journey. Student readiness is not a static trait; rather, it can be intentionally developed over time. Administrators should take pre-emptive measures with course design and focused interventions, like student training that promotes independence and empowerment. Institutional-level policies promoting peer-to-peer cooperation will enable universities to raise the general acceptability of online learning, student involvement, and satisfaction. Instructional support must prioritize clarity and engagement to foster student acceptance of online learning. Institutions can significantly enhance the acceptance of online learning by employing academic and emotional support, integrating technology, and providing comprehensive learning support services. In addition, institutions must constantly build and maintain a solid technology and non-technology support system that includes e-advising, e-tutoring, and mental health counseling for online students. The study advances e-learning practices by reframing student readiness as a dynamic quality that organizations can cultivate with focused instruction and assistance. The results offer practical advice for creating welcoming, stimulating, and encouraging online environments that increase student acceptance of online learning.

Keywords: Online learning acceptance, Readiness for online learning, Peer support, Instructional support, Online learning, Higher education

1. Introduction

Online learning has become a mainstay of higher education in recent years due to technology developments, easy access to the internet, and growing demand for flexible learning choices (Seaman, Allen and Seaman, 2018). Online learning offers a more flexible and customized learning environment, enabling students to engage with course content asynchronously. However, this shift also introduces new challenges, particularly regarding students' acceptance of online learning environments. Educational institutions aiming to ensure student success and enhance learning outcomes in online settings must understand the factors influencing students' acceptance of these platforms (Al-Fraihat, Joy, and Sinclair, 2020). To this end, this study emphasizes three key elements: students' readiness, peer support, and instructional support.

In higher education, online learning must be appropriate, acknowledged, and accepted by students with various capacities and resources (Sharif-Nia, et al., 2024). Students' acceptance of online learning is influenced by their degree of readiness to interact with online learning environments. Among the elements shaping this readiness are personal drive, technological preparedness, and the perceived value of online learning resources. Since

Cite this article: Robinson Jr.,L. 2025. "Exploring Determinants of Online Learning Acceptance: The Role of Readiness, Peer Support, and Instructional Support", *Electronic Journal of e-Learning*, 23(2), pp 130-142, https://doi.org/10.34190/ejel.23.2.4048

online learning heavily relies on students' preparation, motivation, self-discipline, and computer proficiency, their ability to engage effectively in online education varies significantly (Joosten and Cusatis, 2020). This variation in students' engagement underscores the importance of readiness, as adapting to online environments often demands different skills and attitudes than those required in conventional face-to-face settings (Hung, et al., 2010). Students with the necessary skills and internal motivation are more likely to embrace and succeed in online learning environments (Xu, et al., 2023).

In addition to student readiness, peer support is critical in fostering students' acceptance of online learning. Peer support refers to the social, intellectual, and emotional encouragement that students receive from one another throughout their educational journey. Research has shown that peer interactions build a sense of community, increase engagement, and improve learning outcomes in online settings. Establishing and nurturing peer networks in digital environments can help reduce feelings of isolation and create a more interactive and collaborative learning experience (Huang et al., 2023). This is particularly important in online courses, where the lack of real-time social interaction can hinder motivation and involvement (Kebritchi, Lipschuetz, and Santiague, 2017).

Equally important is the role of instructional support in shaping students' perceptions and experiences in online learning. Instructional support includes technical resources facilitating learning and non-technical elements, such as academic assistance, prompt feedback, and a well-structured course design. Lasekan, et al. (2024) highlight the need for support systems (home, school, and peers) to manage the complex nature of the challenges associated with online learning. Furthermore, instructional support helps deal with information overload and perceived technical skill requirements, poor class format, and ambiguous communication, all of which affect perceived learning (Conrad, et al., 2022). Thus, schools must offer technological and non-technical support to promote positive student experiences and the acceptance of online learning environments.

These factors, readiness, peer support, and instructional support, are closely linked to the Technology Acceptance Model (TAM) theoretical framework. TAM posits that perceived usefulness and perceived ease of use are central determinants of technology adoption (Davis, 1989). In the context of online education, the flexibility and customization afforded by asynchronous engagement enhance the perceived usefulness of these platforms. Similarly, students' readiness, encompassing personal enthusiasm, technological competence, and perceived value of online learning resources, plays a crucial role in shaping both the perceived ease of use and the perceived usefulness of online learning environments (Joosten and Cusatis, 2020). Peer support further reinforces these perceptions by mitigating loneliness often experienced in online courses, thus influencing students' willingness to interact (Thongsri, et al., 2021; Sarfraz, et al., 2022). Moreover, instructional support directly impacts students' experiences by reducing barriers to engagement and promoting positive perceptions of online learning platforms (Zhao, Shao, and Su, 2022). Consequently, a comprehensive understanding of how readiness, peer support, and instructional support align to promote acceptance is essential for educational institutions seeking to foster greater adoption and success in online learning environments (Al-Fraihat, Joy and Sinclair, 2020).

Many concerns remain regarding the interactions among the aspects influencing the acceptance of online learning. Studies have been undertaken in higher education environments linking acceptance to issues like achievement outcomes (Szymkowiak and Jeganathan, 2022), economic sustainability (Ahmad, et al., 2023), and process virtualization (Alarabiat, et al., 2024), for example. According to Venkatesh, et al. (2003), a substantial amount of the current published research has focused mostly on technical aspects, such as the alleged simplicity of online learning systems. This trend is still relevant amongst more recent literature (e.g., Salloum, 2018; Baber, 2021; Thongsri, et al., 2021; Szymkowiak and Jeganathan, 2022; Loi, et al., 2023; Halász and Kenesei, 2024; and Sharif-Nia, et al., 2024). Despite this extensive research on online learning acceptance, a critical gap persists in understanding how non-technical factors, such as student readiness, instructional support, and peer interaction, interrelate and influence outcomes. This gap is particularly pressing as higher education increasingly adopts hybrid and fully online models, demanding a more integrated examination beyond technical usability. The broader significance of the study within the e-learning field is that it highlights the need to move beyond technical considerations in online learning environments and address the human and pedagogical dimensions that affect student success and engagement. This research aims to close the identified gap by systematically examining the following research question: How do the nontechnical factors of student readiness, peer support, and instructional support influence students' online acceptance of online learning? By focusing on student readiness, peer support, and instructional support, this research targets human and pedagogical dimensions of online learning that can be crucial to learner success. While individual studies have explored these factors in isolation, few have systematically examined how they interact to influence online learning acceptance. Analyzing these interrelated components will help the research provide insights that will direct institutional policies and support instructors in establishing more successful online learning environments appropriate for a contemporary student population. Doing so provides a more holistic framework for understanding online learning acceptance, ultimately informing more inclusive and effective educational strategies. The remainder of this paper provides a review of the acceptance of online education, student readiness, peer support, and instructional support literature. Next, the theoretical model is empirically tested using structural equation modeling. Finally, a discussion of the model results, implications for higher education institutions, research limitations, and recommended future research is presented.

2. Literature Review

2.1 Students' Online Learning Acceptance

Students' acceptance of online learning reflects their willingness and preparedness to utilize its tools, platforms, and methodologies. Acceptance plays a crucial role in determining the overall success and perceived value of online learning as it helps level the initial environment for students (Loi, et al., 2023). Improving engagement and effectiveness in online learning environments requires understanding the aspects influencing students' acceptance of online learning. Learners' attitudes toward e-learning can be shaped by several personal, societal, and institutional elements, therefore affecting its acceptability (Sharif-Nia, et al., 2024). Robinson (2024) suggested that students consult various professionals for advice on efficient online learning practices. This shows the need for support systems in boosting student confidence and interaction with online learning. Moreover, Zheng, Bender and Lyon (2021) noted that students' acceptance of online learning in higher education environments primarily depends on their positive attitudes towards online learning and instructional strategies.

Strong e-learning systems greatly affect students' willingness to accept online learning, claims Goh and Blake (2021). In line with this, Salloum (2018) underlined that the acceptability and success of e-learning systems in higher education depend on the opinions and attitudes of the students. These results imply that institutions must invest in pedagogical techniques and technology developments to improve student acceptability and involvement in online learning environments. Acceptance of online learning by students is a complicated concept influenced generally by technical, institutional, and human aspects (Chaka and Govender, 2017). Administrators and instructors who know and enable these elements can create more successful online learning environments that satisfy different student needs.

2.2 Students' Readiness for Online Learning

Students' readiness for online learning plays a pivotal role in shaping student perspectives and experiences within online learning spaces. Readiness is a multidimensional construct encompassing various psychological, technological, and behavioral factors, determining students' ability to engage with and accept online learning (Chung, Subramaniam and Dass, 2020). Rajeb, et al. (2023) realize that institutional designs and student-specific factors determine how individuals embrace online learning. The quality of student participation and performance in online learning depends on knowledge of different readiness dimensions (Jayanthi and Rajalakshmi, 2022). Comfort with e-learning explains how comfortable and confident a student is in working with online systems, and self-management focuses on the ability of a student to plan, organize, and control their learning environment. These factors express the requirement for psychological and behavioral readiness. This implies that students with greater self-control are more confident in an online setting.

Hung, et al. (2010) put forth five essential elements of online learning readiness: self-directed learning, motivation for learning, learner control, computer and internet self-efficacy, and online communication self-efficacy. Each component helps a student be competent in navigating, using, and gaining from online learning environments. Particularly important are self-directed learning and motivation since they allow a student to become naturally motivated to interact with the course contents without direct guidance. Learner control is the capacity to adjust to various learning speeds and approaches; technological self-efficacy affects a student's confident interaction with digital tools. Online communication self-efficacy guarantees that students may participate in debates, work with peers, and get help when necessary (Chung, Noor and Mathew, 2020).

Studies confirm that effective online learning outcomes are positively associated with greater degrees of readiness. Those students with a higher degree of readiness, who accounted for taking up course materials, were more self-directed with prior experience using learning technology (Chau, Law and Tang, 2021). Equally, Wei and Chou (2020) established that attitudes among students toward online learning were highly related to how ready they were, which results in smoother transitions to online learning. Research conducted by Ifinedo (2017) and Kirmizi (2015) provided a foundation for this relationship by showing that students with outstanding technological self-efficacy and independent learning techniques had a propensity to embrace and succeed in computer-mediated learning environments. Following these findings, the following hypothesis is formulated.

H1: Students' readiness for online learning positively impacts their acceptance of online learning.

2.3 Peer Support

Peer support is important in preparing students for online learning by providing a sense of belonging, mitigating feelings of isolation, and promoting active engagement. Becoming accustomed to online learning poses challenges, including reduced face-to-face communication and limited real-time access to teachers, which could hinder the pace of the academic performance of students. However, peer support relieves these issues by offering collaborative learning experiences, enabling students to transition to online learning more smoothly. Lee, et al. (2011) emphasized that peer interactions, such as group discussions, group projects, peer teaching, peer tutoring, and peer facilitation, are essential in assisting students to adjust to online learning. Through these interactions, students can exchange knowledge, describe complex concepts, and learn problem-solving skills in collaboration. In addition, these activities allow students to refine their independent learning and critical thinking skills, which are crucial to success in online education.

Besides the cognitive benefits, peer support is important in maintaining students' motivation and emotional well-being. Online learning environments can become isolated, negatively affecting students' persistence and motivation (Muilenburg and Berge, 2005). As students actively mentor one another, they build a more cohesive and inclusive learning environment that counteracts the isolating influence of online learning. Emotional support from peers reduces anxiety, fosters a sense of belonging, and encourages students to remain committed to their studies despite challenges. Students' social and communication skills are critical determinants of their readiness for online learning, and these skills are reinforced through peer interactions (Osman, Mohamad and Mohamad 2021). Effective communication, collaboration, and interpersonal skills obtained in peer-facilitated learning environments make students more confident in participating in virtual discussions, engaging with course materials, and seeking assistance as needed. Furthermore, positive interactions with peers minimize anxiety and support a growth mindset, ultimately increasing students' learning readiness and performance (Muslichah, et al., 2022).

Students who engage in peer-supported learning activities acquire essential digital literacy and technological proficiency, which are fundamental for online education. They become more adept at using learning management systems (LMS), online collaboration tools, and communication platforms, increasing their overall readiness for online learning (Thongsri, et al., 2021; Sarfraz, et al., 2022). The enabling characteristic of peer learning serves to prepare, empower, and make students more confident to handle the dynamics of online learning. Based on these observations, peer support is a crucial antecedent to students' readiness for online learning, facilitating their learning achievement, emotional stability, and social competence. The following hypothesis is proposed:

H2: Peer support positively impacts students' readiness for online learning.

Peer support is a significant element in students' acceptance of online learning because it provides a network of intellectual, social, and emotional support. A supportive online learning community encourages student participation through peer interactions that enhance the learning experience (Zhao, Shao and Su, 2022). The necessity of peer connection is well-documented because the absence of social interaction is among the predominant problems in online learning (Muilenburg and Berge, 2005). When students experience isolation, their motivation to attend online courses diminishes, leading to lower acceptance and satisfaction with online learning systems. Conversely, students who develop meaningful peer relationships will likely perceive online learning as beneficial and become more accepting and adaptive to such learning (Fabriz, Mendzheritskaya and Stehle, 2021).

The impact of peer support spans learning's social interaction, cognitive, and affective aspects. Wei and Chou (2020) concluded that how students perceive teachers and how they perceive their peers determines their overall acceptance and satisfaction with online learning. An effective peer network enhances a feeling of belonging, deters apprehension, and stimulates motivation. Such a feeling of belonging is also crucial in transcending common online learning barriers, such as technical difficulty and loneliness (Halász and Kenesei, 2024). Once students know they can rely on their peers for assistance in comprehending course material or troubleshooting technical issues, their confidence in online learning increases. Furthermore, social influence is crucial in organizing students' attitudes towards online learning. Students' peer interaction is accountable for a collaborative learning culture, which upholds the legitimacy and benefits of learning online (Halász and Kenesei, 2024). Facilitating cooperation through group assignments, peer review, and discussion forums makes it possible to have a collaborative learning environment, where students feel appreciated and respected. Such a collaborative setting, in return, enhances participation and academic achievement, thus making online learning a more attractive and viable option for students.

Theoretically, peer support is aligned with the TAM, indicating that perceived usefulness, ease of use, and social influence are crucial determinants of technology acceptance (Davis, 1989). This alignment suggests that peer support can enhance students' perceptions of the usefulness and ease of online learning, reinforcing the influence of social factors on their acceptance of technology. If students observe their peers effectively using online learning environments and deriving benefits from them, they will learn to appreciate online learning as more acceptable and feasible. Thus, Institutions must prioritize activities that promote peer interaction, such as structured peer mentoring programs, online study groups, and interactive learning interfaces (Szymkowiak and Jeganathan, 2022). Peer support is a precursor to the acceptance of online learning since it fulfills social, cognitive, and emotional requirements. By fostering a feeling of community, removing barriers, and consolidating favorable attitudes towards online learning, peer interactions are vital in enhancing students' intentions to utilize and accept online learning. The discussion above supports the following hypothesis:

H3: Peer support positively impacts students' online learning acceptance.

2.4 Instructional Support

Students' acceptance of online learning depends on their access to effective resources that foster engagement, satisfaction, and perceived learning outcomes, all of which can be supported by instructor guidance. As defined by Lee, et al. (2011), instructional support involves answering questions, clarifying misunderstandings, and providing structured feedback to create an effective online learning environment. This support enhances students' motivation and understanding of the material, emphasizing the need for instructor presence in online classrooms (Lee, et al., 2011). Instructional support allows educators to respond to individual students' needs, measure their progress, and adjust their pedagogical approach accordingly (Ahmad, et al., 2023). Full-scale learning support services, such as advising and study assistance, promote learner engagement and individual and collaborative learning (Zhao, Shao, and Su, 2022).

Instructor feedback and engagement are essential to students' acceptance and success in online education. Research indicates that instruction interactions have a greater impact in virtual settings than in traditional classrooms, which implies that the frequency and quality of instructor interaction are vital to student acceptance and learning engagement (Rajeb, et al., 2023; Lasekan, et al., 2024). Moreover, multiple researchers have identified the strong correlation between instructional presence and learner acceptance of e-learning. Immediate feedback via technology tools, like instant messaging, can increase positive attitudes and acceptance of online learning (Maheshwari, 2021). As part of the quality control of an online learning environment, instructor attitude and interaction are useful (Baber, 2021). Effective interactions between the instructor and student provide more opportunities for learning to take place. Lee (2018) supported that quality instructor-student interaction can lead to more student motivation and inspired learning. According to these studies, learners tend to accept e-learning if they perceive instructional support as being timely, responsive, and personalized. When learners feel that the instructors are individually involved in their learning process, they become more interested in the coursework and participate actively in discussions, leading to greater satisfaction and acceptance of the online mode.

In addition, instructional support can ease the difficulties of self-regulated learning in online settings. In contrast with the more traditional face-to-face instruction, e-learning tends to make students commit themselves to being more responsible learners. Facilitative instruction based on formative feedback could enhance learners' self-efficacy and reduce feelings of isolation (Sun and Rueda, 2012). By addressing learners' academic and emotional requirements, instructional support boosts their general learning process and improves their readiness to adopt online learning as a viable education model. Moreover, instructional support and technological support services contribute to developing effective online learning environments (Lee, 2010). Combining instructional and technological support makes students more likely to achieve greater acceptance and performance in e-learning environments. Rajeb, et al. (2023) emphasize that enhanced instructional support significantly increases students' acceptance and satisfaction with online learning methods. Therefore, the following is hypothesized:

H4: Instructional support positively impacts students' online learning acceptance.

Figure 1 illustrates each of the previously proposed hypotheses and the corresponding conceptual framework.

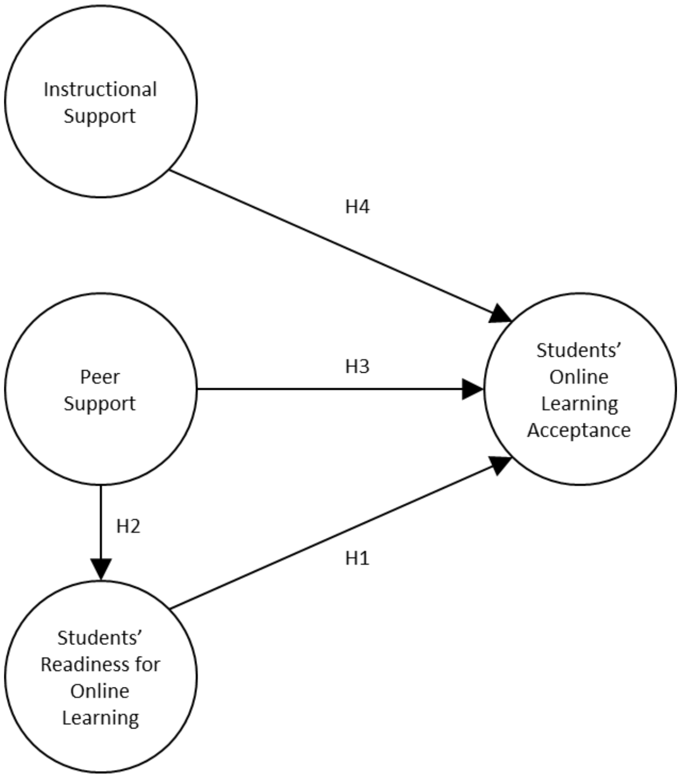

Figure 1: Research Model and Hypotheses

3. Methodology

3.1 Data Collection

The selection approach for this study employed a structured sampling frame that included undergraduate and graduate students currently enrolled in fully online or hybrid courses. Participants were recruited from a four-year university in the southwestern United States. The study adhered to rigorous ethical guidelines, as the Institutional Review Board (IRB) at the researcher's university thoroughly reviewed and approved the research protocol before data collection commenced. The recruitment process was facilitated using the cloud-based SONA Systems platform, which was integrated into the university's research participation system. This platform allowed for efficient participant management and ensured compliance with institutional research guidelines.

All collected data were securely stored in a protected database to maintain strict confidentiality. The research team took extensive measures to anonymize responses, ensuring no personally identifiable information was linked to individual participants. Any findings derived from the study were reported in an aggregated manner to preserve participant privacy. The study yielded 308 usable responses from volunteer participants, who received extra credit as compensation for their involvement. The sample comprised 70% females, 29% males, and 1% individuals with an undeclared gender identity. Additionally, 82% of respondents were undergraduate students, while 18% were graduate students, all representing diverse business disciplines.

3.2 Measurement Scales

The study employed multi-item scales acknowledged and validated within the discipline, distinguished by clearly articulated conceptual frameworks and substantiated by significant evidence of their reliability and validity. All items were selected from established measurement scales; where appropriate, adjustments were implemented to ensure alignment with the study setting. A 7-point Likert scale was utilized to assess responses for all items, with 1 indicating "strongly disagree" and 7 denoting "strongly agree." Compared to other scales, a 7-point Likert scale provides a balanced range of response options, improving measurement sensitivity and permitting more nuance in participants' attitudes without overwhelming respondents as longer scales might. The dependability of each measure was assessed utilizing Cronbach's coefficient alpha. The reliability of each measurement scale was considered acceptable, above the commonly acknowledged threshold of .70 (Hair, et al., 2019). Table 1 shows the reliability of each construct, and Appendix 1 summarizes the measurement scales, their sources, and the items.

Table 1: Reliability of Constructs

Constructs	No. of items	Cronbach's Alpha
Students' Online Learning Acceptance	3	.900
Students' Readiness for Online Learning	6	.881
Peer Support	4	.792
Instructional Support	5	.914

Note. All Cronbach's alphas are significant at the 0.01 level (2-tailed).

4. Results

4.1 Confirmatory Factor Analysis

Brown and Moore (2012) recommended that a confirmatory factor analysis (CFA) be conducted using AMOS Version 29 to evaluate the alignment between observed variables and their respective latent constructs. The measurement model included all construct-related items, which were simultaneously tested with each item constrained to load onto its designated conceptual component. The chi-square test yielded a value of 281.785 with 138 degrees of freedom (df) and a significance level ($p < 0.000$). The normed chi-square value (CMIN/DF = 2.042) was below the critical threshold of 5.0, indicating an acceptable fit. Following established guidelines, additional fit indices were examined to assess model adequacy. Based on criteria from Hair, et al. (2019) and Hu and Bentler (1999), the measurement model demonstrated satisfactory fit, as indicated by the following indices: standardized root-mean-square residual (SRMR) = 0.0521, Goodness-of-Fit Index (GFI) = 0.914, Tucker-Lewis Index (TLI) = 0.951, Comparative Fit Index (CFI) = 0.961, and Normed Fit Index (NFI) = 0.926.

Each variable's average variance extracted (AVE) was examined to assess convergent validity. The AVE for students' online learning acceptance, students' readiness for online learning, peer support, and instructional support exceeded the threshold of 0.50, indicating strong convergent validity (Fornell and Larcker, 1981). Discriminant validity was assessed by comparing the square roots of the AVE values to the maximum squared pairwise correlation (0.583). The results confirmed that the square roots of the AVE values exceeded the correlations between variables, meeting the criteria established by Fornell and Larcker (1981). Furthermore, it was verified that the square roots of the AVE values for each construct were above 0.7 and exceeded the correlations between variables, reinforcing discriminant validity (Fornell and Larcker, 1981). Table 2 displays various characteristics of each measurement scale, including Composite Reliabilities (CR), AVE values, scale correlations, and the square roots of the AVE values.

Table 2: Composite Reliability, Average Variance Extracted, Correlations, and Square Root of AVE

Constructs	CR	AVE	1	2	3	4
Students' Online Learning Acceptance	.903	.700	.837			
Students' Readiness for Online Learning	.875	.539	.549	.734		
Peer Support	.824	.541	.569	.431	.736	
Instructional Support	.916	.687	.516	.367	.548	.829

Note. All correlations are significant at the 0.01 level (2-tailed). Numbers shown on the diagonal denote the square root of the average variance extracted.

4.2 Structural Equation Modeling

Structural equation modeling (SEM) simultaneously examines multiple relationships, accounting for measurement errors and comprehensively evaluating theoretical models (Arbuckle, 2022; Byrne, 2016). This approach is particularly suitable for testing hypotheses (Hair, et al., 2019). Additionally, SEM integrates CFA and path analysis, ensuring that measurement and structural models align with theoretical expectations (Kline, 2023). The analysis demonstrated that the resultant indices, including Chi-square (CMIN = 265.057, df = 141, $p < 0.000$), CMIN/DF (1.880), RMSEA (0.054), SRMR (0.0632), GFI (0.921), TLI (0.959), CFI (0.966), and NFI (0.931), indicated a good fit. An examination was conducted on the structural model's standardized residuals and modification indices, revealing no substantiated rationale for any theoretically significant alterations. These findings align with the recommendations of Hair, et al. (2019) and Hu and Bentler (1999), suggesting that the postulated structural model that underlies the study was deemed appropriate for continued investigation.

All standardized beta estimates had positive values and demonstrated statistical significance. The path coefficients were analyzed to determine if the collected evidence supported the hypotheses. The higher measurement of a student's readiness for online learning was positively related to a higher measurement of the student's online learning acceptance (β = 0.258, t = 4.897). The higher perception of peer support was positively related to a higher measurement of the students' readiness for online learning (β = 0.571, t = 6.451). It was found that a higher perception of peer support was positively related to a higher measurement of the student's online learning acceptance (β = 0.301, t = 3.905). Finally, a higher perception of instructional support was positively related to a higher measurement of the student's online learning acceptance (β = 0.230, t = 4.298). Table 3 presents the beta values, t-values, and tests of hypothesized relationships of the constructs.

Table 3: Results of Hypothesis Testing

	Hypothesis Pathways	β value	t-value	Results
H1	Students' Readiness for Online Learning → Students' Online Learning Acceptance	.258	4.897	Supported
H2	Peer Support → Students' Readiness for Online Learning	.571	6.451	Supported
H3	Peer Support → Students' Online Learning Acceptance	.301	3.905	Supported
H4	Instructional Support → Students' Online Learning Acceptance	.230	4.298	Supported

Note. All t-values are significant at $p < .001$

5. Discussion

The results support a positive relationship between students' online learning readiness and acceptance. A complex concept, readiness consists of psychological, technological, and behavioral elements influencing students' capacity to participate in online learning (Chung, Subramaniam, and Dass, 2020). Thus, increasing acceptance and efficiency depends on the evaluation of readiness. Before or at the beginning of a course, it is essential to measure and foster student readiness by administering structured readiness assessments. More ready students are often confident and creative, improving their capacity to negotiate online learning settings (Yusuf, et al., 2021). Higher readiness, defined by self-regulation, drive, and technological competency, correlates with greater acceptance, as confident students are more inclined to embrace online learning (Ucar and Ugurhan, 2023). Implementing targeted workshops that focus on enhancing digital literacy, time management, and self-regulation skills can provide students with the foundational competencies required for success in engagement with the course. Strong self-efficacy and digital literacy also help students move more naturally into online learning, supporting the connection between acceptance and readiness (Wei and Chou, 2020; Saqr, Al-Somali and Sarhan, 2024). Incorporating confidence-building activities alongside regular, formative feedback can significantly strengthen students' self-efficacy and engagement throughout the learning process. Thus, student acceptance and success depend on encouraging readiness.

The study supports the critical need for peer support in addressing acceptance issues, as an influence on student readiness and acceptance. Peer support becomes crucial given limited face-to-face contact and immediate access to the instructor. Course designers should include mandatory peer interactions, like forums, discussion groups, and peer reviews, to provide structured support to help clarify concepts, exchange knowledge, and solve problems. These interactions are essential for the effectiveness of online education as they foster critical thinking and self-directed learning. Educators should actively promote and integrate structured peer support systems into the design and delivery of online courses. These strategies are effective in enhancing student engagement, motivation, and confidence and are essential for reducing feelings of isolation, a common barrier in online learning environments. Peer mentorship programs can connect experienced or high-performing students with newer learners, providing a valuable avenue for sharing strategies, building digital literacy, and fostering a sense of belonging. Study groups offer regular opportunities for academic collaboration and clarification of course content, encouraging deeper cognitive engagement and reinforcing accountability. Collaborative projects develop teamwork, communication, and problem-solving skills, all while enhancing students' readiness and willingness to accept online learning modalities.

The findings indicate that students who receive peer support exhibit increased engagement, motivation, and cognitive presence in online learning environments. A common issue in online education, isolation, is lessened by peer collaboration, which also helps the learning group to feel belonging (Gao, et al., 2024). Social presence, the view of peers as real and engaging humans in an online environment, helps students participate in higher-order cognitive processes, improving the learning environment (Lee, 2014). This finding is consistent with research demonstrating that knowledge-sharing activities, peer discussions, and group projects contribute to students' academic persistence and success in online learning (Huang, et al., 2023). In order to develop confidence and acceptance in online learning, students should model their behavior on successful classmates. Observing peers effectively manage their online courses motivates students to use the same techniques, enhancing their technological skills. This is consistent with social learning theory, which says people acquire abilities by observing others and mimicking their behavior (Bandura, 1986). Well-run peer mentorship programs give students opportunities for direction, support, and shared experiences, enabling them to accept online learning even more. Programs encouraging interactions, such as debating forums and combined projects, help students to be open to online learning approaches (Mudau and Van den Berg, 2023). By institutionalizing these peer-driven initiatives, administrators and educators can create a supportive and socially rich online environment that encourages active participation, increases persistence, and ultimately improves academic outcomes.

Strong online infrastructure improves student experiences, hence institutional support also matters (Rajeb, et al., 2023). The research emphasizes how instructional support can help students accept online learning. Effective instructional support improves learning, increases involvement, and encourages participation. In online courses, instructors who apply interactive learning techniques give timely comments and use adaptive procedures to raise student motivation (Redmond, et al., 2023). Institutions must provide continuous academic and technical support through virtual office hours and digital help centers to meet student needs. Courses should incorporate multimedia elements, gamification techniques, and a clear, well-organized design to enhance engagement and minimize frustration. Access to these options fosters acceptance of online education and confidence. In online courses, instructional scaffolding, which provides ordered direction while boosting learner autonomy, greatly enhances engagement and performance (Zhu and Bonk, 2022).

Acceptance of online learning also depends on institutional assistance. Universities that invest in technology and non-technology-based tools, educator training, and digital infrastructure raise students' motivation and confidence. Instructors should be trained to effectively use interactive tools, deliver timely feedback, and implement instructional scaffolding strategies that foster student autonomy. Engagement is improved by organized interventions like LMS access, digital literacy courses, and instructor development (Robinson, 2024). Overall, this research deals with filling existing research gaps by reinforcing that online learning readiness is multidimensional and has a bearing on student acceptance of online learning. It integrates previously fragmented literature on readiness, peer support, and institutional factors. It also explicitly links social learning theory to peer support structures and thus provides a theoretical basis for observed behaviors related to online learning. Additionally, it advocates for practical interventions such as readiness assessments, peer mentoring, or buddy systems, and structured digital literacy workshops that have typically been cited in isolation in previous studies.

6. Conclusion

This study significantly impacts e-learning practices by shifting the focus from purely technical aspects of online learning acceptance to a more holistic understanding of the human and pedagogical dimensions involved. Specifically, it highlights how student readiness, peer support, and instructional support are critical nontechnical factors that shape students' willingness to engage in and succeed with online learning. The broader implications of these findings for online education are significant and multifaceted, requiring a holistic approach that not only focuses on content delivery but also actively prepares students, fosters social connections, and supports them throughout their learning journey.

Student readiness promotes success and acceptance in online environments; however, it is not a fixed trait. It can be deliberately cultivated through structured interventions. The study shows that fostering readiness through initiatives that build digital literacy, time management, motivation, and self-directed learning can increase students’ confidence and engagement. Readiness is made up of several components, including technological and social competencies, which enhance a student’s ability to participate effectively in online learning (Chung, Subramaniam, and Dass, 2020). Educators should take proactive steps in course design and implement training programs that promote independence and empowerment. Peer support is emphasized as a powerful mechanism for reducing isolation, enhancing motivation, and building a collaborative learning culture, thereby improving both readiness and acceptance. It drives engagement by increasing sociability, sustaining interest, and fostering investment through cooperative learning. Therefore, policies that promote peer-to-peer cooperation should be implemented at the institutional level and embedded into the core of academic activities, rather than treated as optional. Course design should strategically include peer interaction as a purposeful element to leverage its influence on acceptance and involvement.

Instructional support, including timely feedback, appropriate guidance, and structured communication, is shown to elevate student satisfaction and learning outcomes. For instructional support to foster student acceptance, it must prioritize clarity, engagement, and consistent presence. Institutions can significantly enhance online learning acceptance by integrating academic and emotional support, encouraging instructor-student interaction, and providing comprehensive support services. This includes functional LMS platforms, e-advising, e-tutoring, and mental health counseling. Professional development programs must equip educators with tools to design more interactive online environments. These programs should include training on digital technologies that foster interactivity, methods to encourage peer engagement, and strategies for delivering multimedia feedback effectively (Lasekan et al., 2024). By integrating these nontechnical elements into course design and institutional strategies, this study provides a practical roadmap for improving the effectiveness and inclusivity of e-learning. Its findings support a more student-centered approach, one that advances not only technology adoption but also sustainable academic engagement and success in online learning environments.

7. Limitations and Future Research

This study presents several limitations that should be addressed in future research. One key limitation is sampling bias, as the students were drawn exclusively from a single university. This limited demographic diversity may reduce the generalizability of the findings to students from different institutions or cultural backgrounds. The use of convenience sampling may further compound this issue because the students who volunteer may not accurately represent the broader student population. Additionally, the reliance on self-reported data introduces potential bias through socially desirable responses and memory inaccuracies. Poorly worded or ambiguous survey items may have also contributed to measurement error and respondent uncertainty. To improve the robustness of future studies, research should consider recruiting a more diverse and representative sample using stratified or random sampling methods across multiple universities and regions. Employing mixed-method approaches, such as combining self-reports with direct behavioral observations or academic performance data, can enhance the validity of findings. Moreover, future research should explore longitudinal designs to examine the long-term impact of preparatory programs on students' success and sustained engagement in online learning environments. Investigating the integration of adaptive learning technologies and artificial intelligence tools can also provide insights into how personalized support systems influence student readiness, acceptance, and achievement. These directions will contribute to a more comprehensive understanding of effective practices in online education.

AI Acknowledgment: The authors acknowledge the use of Grammarly (https://grammarly.com/) to help with grammar and punctuation, but did not use it to generate the main content of this paper.

Ethics Statement: All subjects gave informed consent for inclusion before participating in the study. The study was conducted in accordance with the research protocols of the Institutional Review Board (study 24-081) at the researcher's university.

References

Ahmad, S., Mohd Noor, A.S., Alwan, A.A., Gulzar, Y., Khan, W.Z. and Reegu, F.A., 2023. eLearning acceptance and adoption challenges in higher education. *Sustainability*, 15(7), p.6190. https://doi.org/10.3390/su15076190

Alarabiat, A., Hujran, O., Al-Fraihat, D. and Aljaafreh, A., 2024. Understanding Students' Resistance to Continue Using Online Learning. *Education and Information Technologies*, 29(5), pp.5421-5446. https://doi.org/10.1007/s10639-023-12030-x

Al-Fraihat, D., Joy, M. and Sinclair, J., 2020. Evaluating e-learning systems success: An empirical study. *Computers in Human Behavior*, 102, pp.67-86. https://doi.org/10.1016/j.chb.2019.08.004

Arbuckle, J.L., 2022. Amos (Version 29.0) [Computer Program]. IBM SPSS.

Baber, H., 2021. Modelling the acceptance of e-learning during the pandemic of COVID-19-A study of South Korea. *The International Journal of Management Education*, 19(2), p.100503. https://doi.org/10.1016/j.ijme.2021.100503

Bandura, A., 1986. *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Brown, T.A. and Moore, M.T., 2012. Confirmatory factor analysis. In: *Handbook of Structural Equation Modeling*. New York: Guilford Press, pp.361-379.

Byrne, B.M., 2016. *Structural equation modeling with Amos: Basic concepts, applications, and programming*. 3rd ed. Routledge, Taylor and Francis Group. https://doi.org/10.4324/9781315757421

Chaka, J.G. and Govender, I., 2017. Students' perceptions and readiness towards mobile learning in colleges of education: A Nigerian perspective. *South African Journal of Education*, 37(1), pp.1-12. https://doi.org/10.15700/saje.v37n1a1282

Chau, K.Y., Law, K.M. and Tang, Y.M., 2021. Impact of self-directed learning and educational technology readiness on synchronous E-learning. *Journal of Organizational and End User Computing*, 33(6), pp.1-20. https://doi.org/10.4018/JOEUC.20211101.oa26

Chung, E., Noor, N.M. and Mathew, V.N., 2020. Are you ready? An assessment of online learning readiness among university students. *International Journal of Academic Research in Progressive Education and Development*, 9(1), pp.301-317. https://doi.org/10.6007/IJARPED/v9-i4/8177

Chung, E., Subramaniam, G. and Dass, L.C., 2020. Online learning readiness among university students in Malaysia amidst COVID-19. *Asian Journal of University Education*, 16(2), pp.45-58. https://doi.org/10.24191/ajue.v16i2.10294

Conrad, C., Deng, Q., Caron, I., Shkurska, O., Skerrett, P. and Sundararajan, B., 2022. How student perceptions about online learning difficulty influenced their satisfaction during Canada's Covid-19 response. *British Journal of Educational Technology*, 53(3), pp.534-557. https://doi.org/10.1111/bjet.13206

Davis, F. D., 1989. Perceived Usefulness, Perceived Ease Of Use, And User Acceptance. *MIS Quarterly*, 13(September), pp.319–322. https://doi.org/10.2307/249008

Fabriz, S., Mendzheritskaya, J. and Stehle, S., 2021. Impact of synchronous and asynchronous settings of online teaching and learning in higher education on students' learning experience during COVID-19. *Frontiers in Psychology*, 12, p.733554. https://doi.org/10.3389/fpsyg.2021.733554

Fornell, C. and Larcker, D.F., 1981. Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), pp.382-388. https://doi.org/10.1177/002224378101800313

Gao, F., Wang, C., Xie, H. and Hong, J., 2024. Social Interaction and Online Learning Efficiency for Middle School Students: The Mediating Role of Social Presence and Learning Engagement. *Behavioral Sciences*, 14(10), p.896. https://doi.org/10.3390/bs14100896

Goh, P.S.C. and Blake, D., 2021. E-readiness measurement tool: Scale development and validation in a Malaysian higher educational context. *Cogent Education*, 8(1), p.1883829. https://doi.org/10.1080/2331186X.2021.1883829

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E., 2019. *Multivariate Data Analysis.* 8th ed. Hampshire, UK: Cengage Learning EMEA.

Halász, Á. and Kenesei, Z., 2024. Online learning acceptance in higher education—Do we know everything? *Vezetéstudomány-Budapest Management Review*, 55(5), pp.2-19. https://doi.org/10.14267/VEZTUD.2024.05.01

Hu, L. and Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), pp.1-55. https://doi.org/10.1080/10705519909540118

Huang, X., Li, H., Huang, L. and Jiang, T., 2023. Research on the development and innovation of online education based on digital knowledge sharing community. *BMC psychology*, 11(1), p.295. https://doi.org/10.1186/s40359-023-01337-6

Hung, M.L., Chou, C., Chen, C.H. and Own, Z.Y., 2010. Learner readiness for online learning: Scale development and student perceptions. *Computers & Education*, 55(3), pp.1080-1090. https://doi.org/10.1016/j.compedu.2010.05.004

Ifinedo, P., 2017. Examining students' intention to continue using blogs for learning: Perspectives from technology acceptance, motivational, and social-cognitive frameworks. *Computers in Human Behavior*, 72, pp.189-199. https://doi.org/10.1016/j.chb.2016.12.049

Jayanthi, V. and Rajalakshmi, A., 2022. Impact of online learning experience on student satisfaction through student engagement. *International Journal of Health Sciences*, 6(S3), pp.840-848. https://doi.org/10.53730/ijhs.v6nS3.5320

Joosten, T. and Cusatis, R., 2020. Online learning readiness. *American Journal of Distance Education*, 34(3), pp.180-193. https://doi.org/10.1080/08923647.2020.1726167

Kebritchi, M., Lipschuetz, A. and Santiague, L., 2017. Issues and challenges for teaching successful online courses in higher education: A literature review. *Journal of Educational Technology Systems*, 46(1), pp.4-29. https://doi.org/10.1177/0047239516661713

Kirmizi, Ö., 2015. The influence of learner readiness on student satisfaction and academic achievement in an online program at higher education. *Turkish Online Journal of Educational Technology (TOJET)*, 14(1), pp.133-142. https://dx.doi.org/10.5944/openpraxis.12.2.1092

Kline, R.B., 2023. *Principles and practice of structural equation modeling*. New York: Guilford Publications.

Lasekan, O.A., Pachava, V., Godoy Pena, M.T., Golla, S.K. and Raje, M.S., 2024. Investigating factors influencing students' engagement in sustainable online education. *Sustainability*, 16(2), p.689. https://doi.org/10.3390/su16020689

Lee, A.R., 2018. Korean EFL Students' perceptions of instructor interaction in a blended learning class. Senior Editor: Paul Robertson, 122.

Lee, J.W., 2010. Online support service quality, online learning acceptance, and student satisfaction. *The Internet and Higher Education*, 13(4), pp.277-283. https://doi.org/10.1016/j.iheduc.2010.08.002

Lee, S.J., Srinivasan, S., Trail, T., Lewis, D. and Lopez, S., 2011. Examining the relationship among student perception of support, course satisfaction, and learning outcomes in online learning. *The Internet and Higher Education*, 14(3), pp.158-163. https://doi.org/10.1016/j.iheduc.2011.04.001

Lee, S.M., 2014. The relationships between higher order thinking skills, cognitive density, and social presence in online learning. *The Internet and Higher Education*, 21, pp.41-52. https://doi.org/10.1016/j.iheduc.2013.12.002

Loi, C.K., Khamkhien, A., Suki, N.M., Akkakoson, S. and Lee, H.A., 2023. Examining factors influencing students' acceptance of online learning during COVID-19 pandemic: Evidence from Thailand. *Malaysian Journal of Social Sciences and Humanities*, 8(1), pp.e001524-e001524. https://doi.org/10.47405/mjssh.v8i1.1524

Maheshwari, G., 2021. Factors affecting students' intentions to undertake online learning: an empirical study in Vietnam. *Education and information technologies*, 26(6), pp.6629-6649. https://doi.org/10.1007/s10639-021-10465-8
Mudau, P.K. and Van den Berg, G., 2023. Guidelines for supporting a community of inquiry through graded online discussion forums in higher education. *Education Sciences*, 13(9), p.963. https://doi.org/10.3390/educsci13090963
Muilenburg, L.Y. and Berge, Z.L., 2005. Student barriers to online learning: A factor analytic study. *Distance Education*, 26(1), pp.29-48. https://doi.org/10.1080/01587910500081269
Muslichah, I., Nik Abdullah, N., Rozzani, N., Sanusi, S., and Mohd Farid Fernandez, D., 2022, December. Factors Affecting Online Learning Satisfaction among Students in Higher Learning Institutions: an Indonesian Study. In *1st Australian International Conference on Industrial Engineering and Operations Management*, https://doi.org/10.46254/AU01.20220615
Osman, Z., Mohamad, L. and Mohamad, R.K., 2021. Mediating effect of digital readiness on the relationship between online peer collaboration, psychological motivation and online engagement in Malaysian online distance learning higher institutions. *ASEAN Journal of Open & Distance Learning (AJODL)*, 13(2), pp.21-31.
Rajeb, M., Wang, Y., Man, K. and Morett, L.M., 2023. Students' acceptance of online learning in developing nations: Scale development and validation. *Educational Technology Research and Development*, 71(2), pp.767-792. https://doi.org/10.1007/s11423-022-10165-1
Redmond, P., Alexsen, M., Maloney, S., Turner, J., Brown, A., Basson, M., Galligan, L., Lawrence, J. and Henderson, R., 2023. Student perceptions of online engagement. *Online Learning*, 27(1), pp.383-403. https://doi.org/10.24059/olj.v27i1.3320
Robinson, L., Jr., 2024. Leveraging Technical Support Services to Achieve Student Outcomes in Online Learning. *The International Journal of Technologies in Learning*, 31(2): pp.1-20. https://doi.org/10.18848/2327-0144/CGP/v31i02/1-20
Salloum, S.A.S., 2018. Investigating students' acceptance of e-learning system in higher educational environments in the UAE: Applying the extended Technology Acceptance Model (TAM). Technology Acceptance and Adoption Models and Theories View Project Big Data and the Decision Ma. ResearchGate.Net (Issue September).
Saqr, R.R., Al-Somali, S.A. and Sarhan, M.Y., 2024. Exploring the acceptance and user satisfaction of AI-driven e-learning platforms (Blackboard, Moodle, Edmodo, Coursera and edX): an integrated technology model. *Sustainability*, 16(1), p.204. https://doi.org/10.3390/su16010204
Sarfraz, M., Hussain, G., Shahid, M., Riaz, A., Muavia, M., Fahed, Y.S., Azam, F. and Abdullah, M.T., 2022. Medical students' online learning perceptions, online learning readiness, and learning outcomes during COVID-19: the moderating role of teacher's readiness to teach online. In*ternational Journal of Environmental Research and Public Health*, 19(6), p.3520. https://doi.org/10.3390/ijerph19063520
Seaman, J.E., Allen, I.E. and Seaman, J., 2018. Grade change: Tracking online education in the United States. Babson Survey Research Group.
Sharif-Nia, H., Allen, K.A., Arslan, G., Reardon, J., She, L., Ghahrani, N., Rahmatpour, P. and Fomani, F.K., 2024. E-learning acceptance: The mediating role of student computer competency in the relationship between the instructor and the educational content. *Teaching and Learning in Nursing*, 19(1), pp.e5-e10. https://doi.org/10.1016/j.teln.2023.08.001
Sun, J.C.Y. and Rueda, R., 2012. Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology*, 43(2), pp.191-204. https://doi.org/10.1111/j.1467-8535.2010.01157.x
Szymkowiak, A. and Jeganathan, K., 2022. Predicting user acceptance of peer-to-peer e-learning: An extension of the technology acceptance model. *British Journal of Educational Technology*, 53(6), pp.1993-2011. https://doi.org/10.1111/bjet.13229
Thongsri, N., Chootong, C., Tripak, O., Piyawanitsatian, P. and Saengae, R., 2021. Predicting the determinants of online learning adoption during the COVID-19 outbreak: A two-staged hybrid SEM-neural network approach. *Interactive Technology and Smart Education*, 18(3), pp.362-379. https://doi.org/10.1108/ITSE-08-2020-0165
Ucar, H. and Ugurhan, Y.Z.C., 2023. The role of e-learning readiness on self-regulation in open and distance learning. *Turkish Online Journal of Distance Education*, 24(4), pp.146-159. https://doi.org/10.17718/tojde.1231705
Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D., 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), pp.425-478. https://doi.org/10.2307/30036540
Wei, H.C. and Chou, C., 2020. Online learning performance and satisfaction: Do perceptions and readiness matter? *Distance Education*, 41(1), pp.48-69. https://doi.org/10.1080/01587919.2020.1724768
Xu, Z., Zhao, Y., Liew, J., Zhou, X. and Kogut, A., 2023. Synthesizing research evidence on self-regulated learning and academic achievement in online and blended learning environments: A scoping review. *Educational Research Review*, 39, p.100510. https://doi.org/10.1016/j.edurev.2023.100510
Yusuf, F., Mirantika, N., Syamfithriani, T.S., Darmawan, E., and Irawan, D., 2021. Technology readiness and acceptance model as a factor for the use intention of LMS e-Learning in Kuningan University. *Journal of Physics: Conference Series*, 1933. https://doi.org/10.1088/1742-6596/1933/1/012005
Zhao, X., Shao, M. and Su, Y.S., 2022. Effects of online learning support services on university students' learning satisfaction under the impact of COVID-19. *Sustainability*, 14(17), p.10699. https://doi.org/10.3390/su141710699
Zheng, M., Bender, D. and Lyon, C., 2021. Online learning during COVID-19 produced equivalent or better student course performance as compared with pre-pandemic: empirical evidence from a school-wide comparative study. *BMC Medical Education*, 21(495). https://doi.org/10.1186/s12909-021-02909-z

Zhu, M. and Bonk, C.J., 2022. Guidelines and strategies for fostering and enhancing self-directed online learning. Open learning: *The Journal of Open, Distance and E-Learning*, pp.1-17. https://doi.org/10.1080/02680513.2022.2141105

Appendix 1: Scale Items and Factor Loadings

Construct/Item (Source)	**Loadings**
Students' Online Learning Acceptance (Lee, 2010)	
If I need to study for advanced degrees (programs), I expect to use my school's online learning system.	.690
If asked, I would likely recommend my school's online learning system as an ideal learning platform.	.905
For future advanced degrees (programs/certificates), I would probably use my school's online learning system.	.834
Overall, I am satisfied with my school's online learning system.	.847
Students' Readiness for Online Learning (Hung, et al., 2010)	
I manage my time well.	.758
I set up my learning goals	.742
I have higher expectations for my learning performance.	.762
I can direct my own learning progress.	.759
I have the motivation to learn.	.719
I improve from my mistakes.	.654
Peer Support (Lee, et al., 2011)	
In my online courses, I enjoyed the group discussions.	.701
In my online courses, there were many opportunities to interact with peers.	.836
In my online courses, I felt that I was respected by other students.	.704
In my online courses, students were willing to provide help to other students.	.682
Instructional Support (Lee, et al., 2011)	
In my online courses, I felt that I could ask any questions regarding the course materials to the instructor.	.803
In my online courses, there were appropriate ways of communicating with the instructor.	.812
In my online courses, I felt that the instructor was easily accessible.	.929
In my online courses, the instructor encouraged students to be successful in this course.	.705
In my online courses, the instructor responded to students' questions in a timely manner.	.877

Note. All measures employed a 1–7 Likert-type scale.

AI-Based Analysis of Student Frustration: Speech and Facial Expression Recognition

Akniyet Tokhtarov[1], Nurgul Toxanbayeva[2], Meirambala Seisekenova[3], Talgat Baidildinov[4], Dametken Baigozhanova[5] and Asylbek Abden[6]

[1]Department of Pedagogy and Psychology, I. Zhansugurov Zhetysu University, Taldykorgan, Kazakhstan

[2]Department Of General and Implied Psychology, Al-Farabi Kazakh National University, Almaty, Kazakhstan

[3]Department Of Economics and Service, I. Zhansugurov Zhetysu University, Taldykorgan, Kazakhstan

[4]Department of Special Pedagogy, Abay Kazakh National Pedagogical University, Almaty, Kazakhstan

[5]Higher School of Information Technology and Engineering, Astana International University, Astana, Kazakhstan

[6]Department of Physical Culture and Primary Military Training, I. Zhansugurov Zhetysu University, Taldykorgan, Kazakhstan

akniyettokhtarov@outlook.com
toxanbayevanurgul@outlook.com
meirambalaseisekenova@gmail.com (corresponding author)
baidildinovt@outlook.com
dbaigozhanova@gmail.com
asylbekabden.edu@outlook.com

https://doi.org/10.34190/ejel.23.2.4043

Abstract: Frustration is a key affective state that affects student engagement and learning outcomes. While mild frustration can promote persistence in problem-solving, prolonged frustration often leads to disengagement and reduced academic performance. In traditional learning environments, instructors rely on facial expressions, vocal cues, and behavioral indicators to identify frustration and provide timely support. Such monitoring becomes impractical in large or digital classrooms. Artificial intelligence (AI)-based emotion recognition offers a scalable solution by automatically detecting frustration through facial and speech analysis, enabling adaptive interventions in real time. This study proposes a multimodal AI system that integrates facial expression recognition using a Convolutional Neural Network (CNN) and speech emotion recognition with a Transformer-based model. The system uses attention-based feature fusion to improve accuracy by weighting the more informative modalities. The model was trained on benchmark datasets, including DAiSEE, IEMOCAP, and RAVDESS, and evaluated in a real-world study involving 160 Kazakhstani university students in online and in-person learning sessions. AI-generated predictions were compared with instructor assessments to validate the system's performance. Results indicate that the multimodal system outperforms unimodal approaches, achieving 85% accuracy, 83% precision, and 86% recall on benchmark data, with 84% accuracy and precision in real-world conditions. Comparative analysis reveals that speech-based cues are more informative than facial expressions, particularly when frustration is masked or internalized. The system is less effective at detecting subtle frustration, highlighting the need for greater contextual sensitivity. Although limitations remain, the results demonstrate the system's potential for scalable implementation in classrooms and online platforms. These findings support the integration of AI-driven frustration detection into adaptive learning platforms to help educators identify students at risk of disengagement. By enabling timely intervention and support, such tools can contribute to more responsive and inclusive educational environments. Future research should explore cultural variation in emotional expression and long-term effects on learning outcomes.

Keywords: Frustration detection, Emotion recognition, Multimodal learning, Facial analysis, Speech emotion recognition, AI in education

1. Introduction

Frustration in learning arises when students face challenges that hinder knowledge acquisition. It can result from unresolved confusion, complex tasks, or inadequate instructional support, influencing academic engagement and performance (Baker et al., 2025). Minor frustration may promote persistence, but prolonged frustration often leads to disengagement and stress (Rahman et al., 2024). A survey involving 22,983 Chinese college students found that 59.9% experienced academic burnout, which can be associated with frustration, particularly in high-pressure environments (Liu et al., 2023). Therefore, the recognition and mitigation of frustration are

Cite this article: Tokhtarov, A. et al. 2025. "AI-Based Analysis of Student Frustration: Speech and Facial Expression Recognition", *Electronic Journal of e-Learning*, 23(2), pp 143-157, https://doi.org/10.34190/ejel.23.2.4043

essential to support student motivation and performance. Additionally, understanding student frustration can help educators foster a psychologically supportive learning environment, allowing early interventions to reduce emotional strain and prevent dropout.

Traditional classroom instructors can detect frustration through facial expressions, voice tone, and behavioral cues, which enables timely intervention. In digital learning environments, frustration is harder to identify and address. Studies show that approximately 59% of students' frustration with e-texts is linked to extraneous cognitive load, 19% stems from technological difficulties, and 28% from curriculum-related issues (Novak, McDaniel and Li, 2023). If left unresolved, these frustrations can adversely affect motivation and learning outcomes, underscoring the need for more effective digital support systems.

In online learning environments and large classroom settings, personalized support is limited. Artificial intelligence (AI)-based automatic emotion recognition can bridge this gap by detecting frustration in real time from students' affective cues (Henderson et al., 2021; Corza-Vargas et al., 2024). Beyond academic interventions, automated frustration recognition can serve as an early-warning system for educators to identify students who may need additional psychological consultation. By tracking emotional trends, teachers and counselors can proactively provide emotional support or recommend mental health resources as needed.

Emotion recognition technology typically utilizes facial expressions and voice signals. Computer vision is employed to track facial movements, while speech analysis examines vocal features such as pitch and intensity (Malekshahi, Kheyridoost and Fatemi, 2024). However, single-modal approaches have notable limitations. Facial expressions may be ambiguous, and speech analysis can be unreliable in noisy conditions (Agung, Rifai and Wijayanto, 2024). Frustration can also be masked in one modality while being evident in another. Moreover, models trained on controlled datasets may not generalize well to real-world educational settings. A multimodal framework that integrates facial and speech cues is therefore required to improve accuracy and robustness (Henderson et al., 2021). Most prior studies in emotion recognition have concentrated on engagement and boredom, often relying on single-modal data such as facial expressions or interaction logs, which do not adequately capture the complexity of frustration (Moon et al., 2022). Speech-based emotion recognition remains underexplored in this domain (Qian and Han, 2022). This study seeks to address these gaps by developing a multimodal approach for more accurate frustration detection.

The research involves the development and validation of an AI-based system for detecting student frustration through facial expression and speech emotion analysis. The research consists of two phases: (1) developing a hybrid model that integrates a CNN for facial analysis and a recurrent neural networks (RNN) or a transformer-based model for speech recognition, and (2) empirical validation through controlled learning experiments.

This study addresses the following research questions (RQ):

RQ1: Can a multimodal AI model significantly improve frustration detection compared to single-modal approaches?

RQ2: How well do automated predictions align with human (instructor) assessments of frustration in real learning scenarios?

RQ3: What are the practical benefits and challenges of implementing such technology in educational settings?

RQ4: How can machine-driven frustration detection assist educators in identifying students who may need additional psychological support or consultation?

The findings of this study contribute to the growing field of affective computing by presenting a novel multimodal approach for frustration detection in education. Scientifically, the research evaluates the effectiveness of integrating facial and speech cues for emotion recognition and identifies sources of classification errors. Comparisons between different model variations (e.g., face-only vs. voice-only vs. multimodal) provide insights into the added value of each modality. Practically, the study offers a prototype system that could be incorporated into e-learning platforms or intelligent tutoring systems to enhance student support. Additionally, the discussion on ethical implications, such as obtaining student consent, avoiding biases, and protecting privacy, provides clear guidance for responsible use of AI in education. Moreover, intelligent frustration tracking can help educational institutions improve their psychological climate by identifying patterns of emotional distress among students. By integrating frustration detection with psychological counseling services, schools can provide targeted support, ensuring that students receive the help they need before frustration negatively impacts their well-being and academic performance.

By bridging the gap between theoretical advancements in artificial intelligence and real-world applications, this study lays the foundation for intelligent frustration detection systems that foster a psychologically supportive learning environment, improve student engagement, and enhance educational outcomes.

2. Literature Review

2.1 AI in Education: Advances in Multimodal Emotion Recognition

The integration of artificial intelligence in education has led to numerous emotion-aware systems capable of detecting and analyzing students' affective states, including frustration (Bustos-López et al., 2022). AI-driven emotion recognition methodologies predominantly leverage computer vision and speech analysis, utilizing CNNs for facial expression classification and RNNs or transformers for speech-based affect detection (Abbaschian, Sierra-Sosa and Elmaghraby, 2021; Wang, 2022). Despite advances, real-world use faces challenges like expression variability, cultural differences, and ethical concerns pertaining to data privacy and surveillance (Banzon, Beever and Taub, 2024).

While machine-driven affect recognition offers considerable advantages over conventional self-reporting mechanisms or interaction log analyses, its practical efficacy is constrained by dataset limitations and generalization challenges (Moon et al., 2022). Many widely utilized datasets, such as DAiSEE (Gupta et al., 2022), provide valuable benchmark resources but lack ecological validity due to controlled conditions (Aguilera, Mellado and Rojas, 2023). Similarly, speech-based models often struggle with spontaneous discourse, regional accent variations, and ambient noise present in authentic classroom interactions (Song et al., 2021). Addressing these constraints necessitates the development of more robust, adaptable models trained on heterogeneous datasets reflective of real-world learning environments.

2.2 Frustration in Learning: Cognitive and Behavioral Correlates

Frustration in education involves cognitive load, emotional distress, and behavioral disengagement when students face academic obstacles (Pekrun and Marsh, 2022). It manifests as an emotional response to perceived obstacles in learning (Baker et al., 2025). Frustration arises from diverse sources, including unclear instructional guidance, excessive task complexity, and delayed instructor feedback, all of which influence learning outcomes (Henderson et al., 2021). Moderate frustration levels may foster problem-solving skills, sustained frustration is correlated with increased stress, academic disengagement, and attrition (Graesser and D'Mello, 2012).

Empirical research has demonstrated that frustration can be conveyed through a combination of facial, vocal, and behavioral indicators, including tense expressions, strained vocal tone, and task disengagement (Moon et al., 2022; Shou et al., 2024). However, AI-based frustration detection models frequently exhibit classification errors due to emotional similarity with states such as confusion and boredom. Confusion, for instance, is a precursor to frustration but does not inherently signal emotional distress, thereby complicating automated classification (Rahman et al., 2024). Moreover, frustration expression is context-dependent, influenced by task complexity, individual learning history, and cultural norms, necessitating intelligent recognition systems capable of integrating contextual variables alongside multimodal affective cues (Henderson et al., 2021).

2.3 Comparative Evaluation of AI-Based Frustration Detection Models

AI-based frustration detection methodologies typically follow unimodal or multimodal analytical frameworks. Unimodal models, such as facial expression (Solanki and Mandal, 2022) or speech-based systems (Song et al., 2021), often perform poorly due to limited input. CNN-based facial models, though accurate in benchmarks, are sensitive to lighting, occlusion, and expression variability (Pordoy et al., 2024; Pham et al., 2023). Similarly, speech emotion recognition models, while effective in controlled environments, exhibit performance degradation in real-time applications due to background noise and spontaneous linguistic variations (Villegas-Ch et al., 2023).

Multimodal fusion models have demonstrated superior performance by integrating facial and vocal features, resulting in higher frustration classification accuracy (Moon et al., 2022). Such models often use feature-level fusion, combining visual and vocal embeddings to improve robustness. Despite their advantages, multimodal approaches face challenges related to computational cost, real-time deployment, and dataset bias. Models trained on narrow datasets often generalize poorly across student demographics, highlighting the need for broader data sources (Bustos-López et al., 2022). Inconsistent annotation practices make it harder to reach consensus on what qualifies as frustration in different educational settings. Table 1 summarizes key models and their methodological strengths and limitations.

Table 1: Comparative Analysis of AI-Based Frustration Detection Approaches

Study	Modalities	Model Approach	Key Findings	Limitations
Solanki and Mandal (2022)	Facial video	Custom CNN + ANN on DAiSEE	86.6% accuracy for frustration detection	Limited to visual cues; lacks contextual integration
Moon et al. (2022)	Facial video + interaction logs	Supervised multimodal fusion on custom dataset	10% performance improvement over unimodal models	Small sample (31 students); tested in controlled settings
Song et al. (2021)	Speech audio	Wide ResNet on spectrograms from game-play corpus	Enhanced classification accuracy over baseline CNN	Absence of visual cues; non-educational domain focus
Rahman et al. (2024)	Facial video + speech	Deep learning fusion model on EmoDetect	Improved robustness in online learning	Ethical concerns regarding student privacy

These findings confirm the advantages of multimodal fusion, despite challenges in dataset diversity, real-world use, and ethical concerns (Mamieva et al., 2023). Multimodal methods consistently outperform single-modality approaches, supporting the use of combining facial and contextual data. However, small samples and narrow models scopes still limit generalizability. For example, Solanki and Mandal (2022) reported high accuracy using detailed facial features, though retraining is likely required for other contexts.

Recent studies demonstrate the growing relevance of multimodal emotion recognition (MER) for educational contexts where accurate detection of affective states such as frustration is critical. Transformer-based architectures offer state-of-the-art performance by capturing complex dependencies between modalities (Lian et al., 2023). Dual-attention mechanisms enhance cross-modal alignment, particularly in speech and facial inputs (Zaidi, Latif and Qadir, 2024). Models using tensor product fusion and transformer backbones have surpassed 93 percent accuracy in recognizing student emotions during learning tasks (Xiang et al., 2024). Body gesture data has also proven valuable, with trimodal systems achieving high accuracy by integrating facial expressions, speech, and posture (Yan et al., 2024). Graph-based reasoning networks like Emotion-LLaMA support fine-grained emotion interpretation and contextual reasoning (Cheng et al., 2024). Systematic reviews emphasize the need for broader dataset diversity and ethical deployment in classrooms (Ahmed, Al Aghbari and Girija, 2023; Khare et al., 2024). Overall, these advancements confirm the potential of MER technologies to support emotionally responsive learning environments when implemented with consideration for practical, cultural, and ethical constraints.

2.4 Ethical Considerations in AI-Based Emotion Recognition

The deployment of automated emotion recognition in education requires close examination of ethical implications, particularly concerning privacy and algorithmic bias. These systems rely on sensitive biometric data, such as facial images and voice recordings, raising concerns about data security and informed consent (Mattioli and Cabitza, 2024). If unregulated, such tools may create a surveillance-oriented environment, where students modify their behavior due to constant monitoring, potentially undermining pedagogical efficacy (Rhue, 2018). Ethical technology deployment mandates transparency, student autonomy in data sharing, and localized data processing to mitigate privacy risks (Mattioli and Cabitza, 2024). Recent studies also emphasize the need for transparent and secure educational technology infrastructures (Sakhipov et al., 2022).

Algorithmic bias is another key concern. Emotion classifiers often vary in accuracy across demographic groups, leading to differential classification outcomes (Rhue, 2018). Cultural differences in emotional expression further complicate generalizability, calling for fairness-aware design and diversified training datasets (Corza-Vargas et al., 2024). Interdisciplinary research is essential to ensure ethical alignment with pedagogical and legal standards.

Future research should focus on expanding dataset diversity, refining fusion models, and enabling adaptive real-time learning environments. Longitudinal studies are also needed to assess the long-term impact of frustration detection on motivation and academic resilience (Baker et al., 2025). Additionally, incorporating context, such as task difficulty and prior performance, may enhance classification and intervention accuracy (Moon et al., 2022). With stronger technical design and ethical safeguards, emotion-aware AI can become a transformative tool for supporting student learning while preserving privacy and autonomy.

Overall, the literature affirms the promise of multimodal emotion detection in education. Yet, issues remain around generalizability, data inclusivity, and ethical implementation. Comparative analyses reveal that while multimodal models outperform unimodal ones, their success depends on diverse data, adaptable design, and bias mitigation. This study addresses these challenges by proposing a robust frustration detection framework, validating it in authentic settings, and offering ethical guidelines for fair adoption.

3. Materials and Methods

3.1 Data Sources and Preprocessing

This study investigates AI-based frustration detection by developing and evaluating a multimodal recognition system. The research comprises two primary phases: (1) the development of an AI model integrating facial expression and speech emotion recognition and (2) empirical validation through controlled learning experiments.

The model was trained using benchmark datasets, including DAiSEE (frustration-labeled video data), IEMOCAP (emotionally expressive speech), and RAVDESS (acted multimodal emotional expressions). To assess real-world applicability, an experimental study was conducted with 160 university students in structured learning scenarios designed to elicit frustration through technical disruptions, complex problem-solving tasks, and delayed instructor feedback. AI-based predictions were systematically compared with instructor evaluations to assess classification accuracy, practical feasibility, and ethical considerations related to privacy and bias. A detailed breakdown of the datasets used is presented in Table 2.

Table 2: Summary of Datasets Used for Model Training

Dataset	Samples	Participants	Emotion Labels	Modality
DAiSEE (Gupta et al., 2022)	9,068 videos	112 users	Boredom, Confusion, Engagement, Frustration (4 intensity levels)	Video
IEMOCAP (Busso et al., 2008)	302 dialogues	10 speakers (5 pairs)	Angry, Excited, Fear, Sad, Surprised, Frustrated, Happy, Disappointed, Neutral	Audio-Video
RAVDESS (Livingstone and Russo, 2018)	7,356 audio-visual files	24 actors (12 female, 12 male)	Calm, Happy, Sad, Angry, Fearful, Surprise, Disgust (speech) + Song emotions	Audio-Video, Audio-only, Video-only

3.2 AI Model Architecture

The multimodal AI architecture, illustrated in Figure 1, integrates two primary branches: a convolutional neural network for facial expression analysis and a transformer-based model for speech emotion recognition.

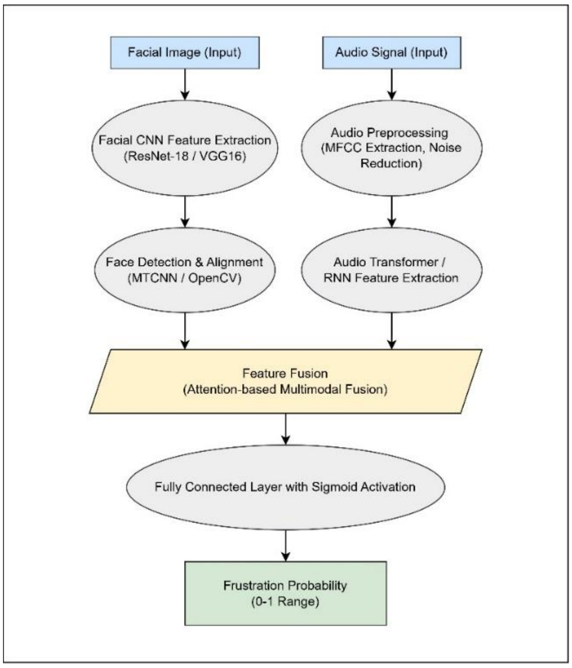

Figure 1: Multimodal AI Architecture for Frustration Detection, Illustrating the Facial CNN, Audio Transformer, and Fusion Layers

The CNN, based on ResNet-18 or VGG16, extracts both low- and high-level features relevant to frustration detection, such as textures, furrowed brows, and narrowed eyes. A feature pyramid module enhances the model's ability to capture fine-grained facial expressions (Mamieva et al., 2023). Instead of direct classification, the CNN generates an embedding vector, which is smoothed over time to reduce sensitivity to brief fluctuations. For speech processing, the model analyzes Mel-frequency cepstral coefficients (MFCCs) to extract frustration-related vocal cues. Both LSTM and Transformer architectures were considered, with the Transformer outperforming LSTM in capturing early vocal indicators of frustration.

Feature fusion was implemented using an attention-weighted strategy, dynamically prioritizing facial or vocal cues depending on their informativeness (Wang et al., 2023; Zaidi, Latif and Qadir, 2023). The final classification layer applies a sigmoid activation function, using binary cross-entropy loss for optimization. Dropout and L2 regularization were included to prevent overfitting.

The CNN and speech models were pre-trained separately before being combined for joint fine-tuning. The fusion mechanism adapts dynamically, giving priority to facial expressions when vocal signals are unclear, and vice versa. This approach significantly improves accuracy over unimodal models (Moon et al., 2022). To minimize false positives, frustration is detected only when both modalities indicate it, or when at least one parameter exceeds a critical level. The architecture is depicted in Figure 1, illustrating the integration of facial and vocal modalities within the frustration detection system.

3.3 Experimental Setup

The proposed model was evaluated through two main experimental phases: (1) offline evaluation on curated datasets, using train/validation/test splits to measure baseline performance, and (2) real-time experimental testing with student volunteers in real learning scenarios to assess real-world effectiveness and compare AI-based frustration detection with human observations. The overall structure of these experimental phases is illustrated in Figure 2.

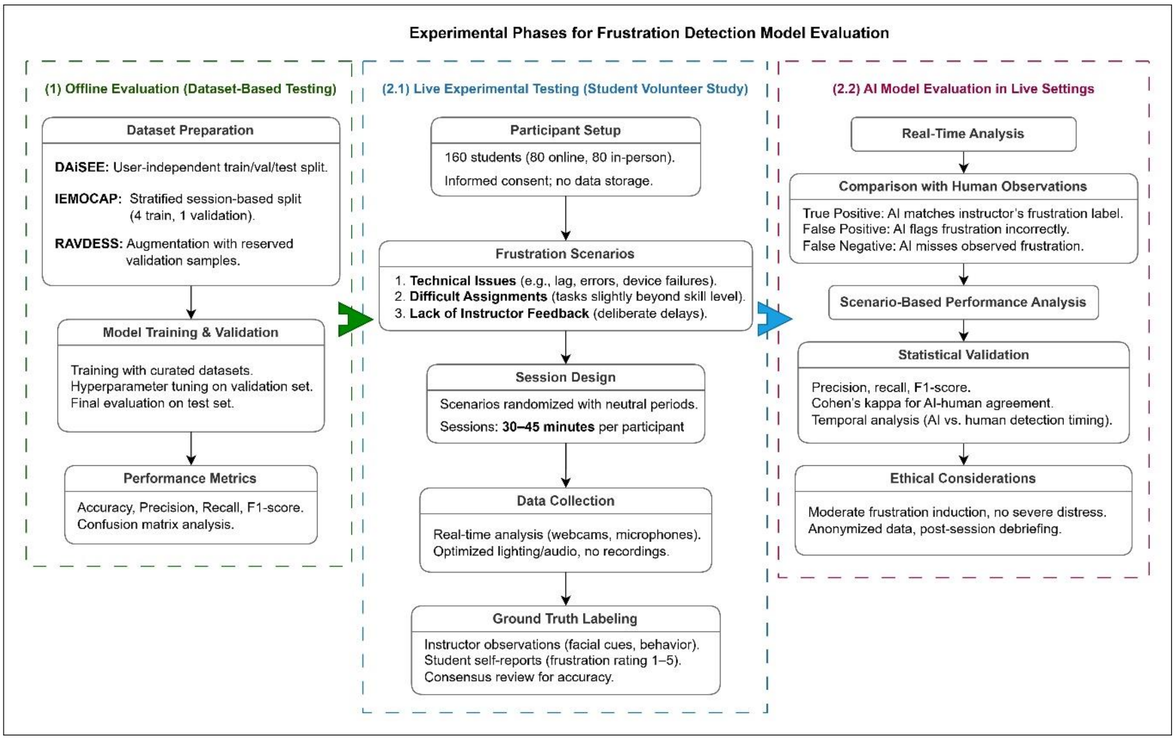

Figure 2: Experimental Phases for Frustration Detection Model Evaluation

This research followed the official DAiSEE partitions: the training set for model training, the validation set for hyperparameter tuning, and the test set for final evaluation. DAiSEE ensures a user-independent split (no overlap between train/val/test), which helps reduce overfitting. For the IEMOCAP audio, following the approach of Busso et al. (2008), a stratified split was conducted based on the five dialogue sessions. Four sessions were used for training and one for validation, rotating for cross-validation due to the small speaker pool. RAVDESS, following the framework outlined by Livingstone and Russo (2018), primarily served to augment training. A subset was used for mini-validation to monitor overfitting.

The evaluation employed standard classification metrics: Accuracy, Precision (*true positives / [true positives + false positives]*), Recall (*true positives / [true positives + false negatives]*), and F1-score (harmonic mean of precision and recall). These metrics were calculated on both held-out dataset partitions and subsequently validated with volunteer testing data. Confusion matrices helped analyze frequent misclassifications (e.g., confusion vs. frustration).

A total of 160 university students joined controlled learning sessions. Participants provided informed consent, and no media were stored. Half participated online via personal computers and webcams, and half attended in-person classes with cameras and microphones, addressing different practical challenges related to equipment setup.

Frustration scenarios were based on common triggers from prior research: (1) technical issues (e.g., slow loading, errors, or hardware malfunctions); (2) difficult assignments, slightly beyond skill levels to induce productive struggle; and (3) lack of instructor feedback, where instructor responses were intentionally delayed, creating feelings of being unsupported. Each participant faced randomized frustration scenarios alongside baseline tasks. Sessions lasted approximately 30–45 minutes.

Student facial and vocal data were analyzed in real time via webcams (online) or classroom cameras and microphones (in-person), without storing recordings. Lighting and audio conditions were adjusted for realistic settings. Frustration instances were identified using scenario timestamps, instructor observations, and student self-reports. Instructors or trained researchers monitored facial expressions, vocal reactions, and behavioral cues (e.g., sighs, frowns, disengagement). Participants rated their frustration on a 1–5 scale, refining observer assessments. Discrepancies were resolved through second rater consensus to ensure labeling accuracy.

After obtaining informed consent, the model processed facial and speech data in real time using a sliding-window approach, generating frustration probabilities flagged when exceeding optimized thresholds. Detected instances, logged with timestamps, were compared to human observations to assess true positives, false positives, and false negatives. Performance metrics such as precision, recall, accuracy, and F1-score were used to evaluate scenario-specific strengths and weaknesses (Table 3). Ethical guidelines ensured minimal distress, with real-time processing conducted anonymously without data retention. Post-session debriefings confirmed participant well-being. Statistical analyses, including Cohen's kappa, measured alignment between AI predictions and human labels, validating the model's effectiveness and identifying areas for improvement.

3.4 Ethical Considerations

Ethical principles guided all stages of this study, with particular focus on privacy, informed consent, and bias mitigation. To protect participants, no video, photo, or audio recordings were stored during testing; only real-time outputs were analyzed. Results were anonymized, and all visuals used in analysis were either blurred or abstracted. The system is designed to function without storing raw data, allowing real-time, local processing on user devices. These measures ensure compliance with the General Data Protection Regulation (GDPR), including principles of data minimization and informed processing.

All participants provided informed consent after being clearly briefed on the study's goals, data use, and their right to withdraw. It was explicitly stated that the AI was a research tool and not a diagnostic system, reducing any psychological pressure. Bias mitigation efforts included the use of diverse datasets (e.g., DAiSEE, IEMOCAP, RAVDESS) and a participant pool representing multiple genders and ethnicities. A multimodal fusion approach helped reduce bias from any single input source.

The system is intended to support learning, not monitor or penalize students. Its outputs are meant to prompt supportive interventions, not judgments. Ethical safeguards were applied throughout to prioritize student well-being and ensure the system remains a responsible and learner-centered tool.

4. Results and Discussion

4.1 Performance Analysis of AI-Based Frustration Detection

4.1.1 Phase 1: Results of offline evaluation (dataset-based testing)

As shown in Table 3, the multimodal model outperformed the single-modality baselines. Precision (83%) and recall (86%) show the model detects frustration accurately without excessive false positives. This balance suggests the model is both effective at identifying frustration and avoiding excessive false positives – a desirable trait for practical use. The model's success is attributed to the complementary strengths of facial and voice signals: when one modality failed, the other often compensated. For instance, a sample with a neutral facial expression but frustrated vocal tone was correctly flagged by the audio model. Conversely, another case with a calm voice but a scowling face was accurately classified by the vision model. The attention-based fusion mechanism dynamically prioritized the more informative input, improving robustness. Benchmarks from prior studies suggest that this F1-score (0.85) is competitive. While direct comparisons are difficult due to dataset differences, the performance is higher than some prior multimodal models, particularly in binary frustration classification.

Table 3: Performance of Unimodal vs. Multimodal Models on Test Data

Model Variant	Accuracy	Frustration		
		Precision	Recall	F1-score
Facial CNN (vision only)	75%	0.78	0.70	0.74
Audio Transformer (audio only)	78%	0.80	0.75	0.77
Multimodal CNN+Audio (fusion)	**85%**	**0.83**	**0.86**	**0.85**

As summarized in Table 3, the multimodal fusion achieved a notable F1-score improvement (+11%) over unimodal models, demonstrating its practical value in capturing frustration that might be missed when relying solely on facial or vocal cues.

4.1.2 *Phase 2: Results of real-time experimental testing*

The real-world evaluation tested the model on 160 students across 216 frustration-inducing scenarios. Instructor observations confirmed 207 actual frustration instances, while the AI flagged 186 instances (Table 4), of which 178 were true positives, 8 were false positives, and 30 were missed detections. Additionally, the model correctly identified 32 cases where students were not experiencing frustration (true negatives), distinguishing them from similar emotions like confusion or concentration. The model achieved 84% precision and 86% recall, closely aligning with its offline test performance. This suggests it generalizes well, though some cases of internalized frustration were missed, and confusion was occasionally misclassified as frustration.

To further analyze the model's classification performance, Figure 3 presents the confusion matrix for frustration detection. The matrix illustrates true and false classifications in the test dataset. The true label represents the actual frustration state as determined by human observers, while the predicted label indicates the AI's classification output.

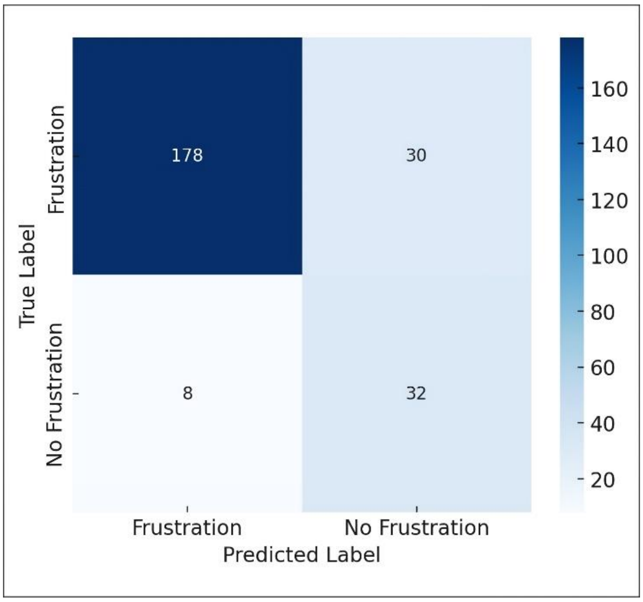

Figure 3: Confusion Matrix for Frustration Detection

Table 4 shows that frustration induced by delayed feedback was the most difficult for students, and also yielded the highest detection rate (87%). However, confusion during challenging tasks led to more false positives, suggesting a need for better differentiation.

Table 4: Frustration Detection Results in Different Learning Scenarios (N=160 students)

Frustration Scenario	Instances (students) with frustration (per instructor)	Detected Cases of Frustration	Detection Rate (Recall)	Noted False Alarms
Technical Issues (online)	40/80 (online group only, N=80 online students, 40 showed frustration)	36 out of 40 actual cases	90%	1 (in 2 cases model thought frustration during minor lag that student shrugged off)
Difficult Assignment	56/160 (some students enjoyed challenge)	46 out of 56 cases	~82.14%	3 (flagged students as frustrated, but they reported only mild confusion)
Lack of Feedback	120/160 (most students found being ignored frustrating)	104 out of 120 cases	~86.7%	4 (brief frustration was flagged in cases where students were only slightly annoyed)

(Note: Each scenario was conducted for each student. "Instances with frustration" is how many students actually felt frustrated in that scenario according to observation/self-report. "Detected Cases" is how many of those instances the system successfully flagged. False alarms indicate cases where frustration was flagged, but observers did not confirm its presence.)

Tables 3 and 5 show the model's high precision (83% on test data, 84% in experiments), meaning most flagged frustration instances were accurate. However, optimizing the detection threshold is crucial because lowering it boosts recall but risks more false alerts, which may be impractical in a classroom. Qualitative observations revealed a reliance on overt signals. Vocal expressions such as sighs triggered instant detection, while silent frustration took longer to register due to temporal smoothing. Masked emotions, such as polite smiles covering irritation, often went undetected, reflecting a limitation shared by human observers. Interestingly, participants who quickly shifted from frustration to focus were sometimes initially misclassified as frustrated, but the probability decreased as their demeanor stabilized. This suggests the system captured momentary states rather than persistent frustration. Overall, the model effectively detected frustration when clear cues were present but struggled with internalized frustration and confusion misclassification. Further improvements may include context-aware features (e.g., task difficulty) and adaptive sensitivity based on individual patterns.

Figure 4 shows that AI-based frustration detection closely matches instructor evaluations, particularly for technical issues (90%) and lack of feedback (87%). While the model slightly underperforms in recognizing frustration from difficult assignments (82% vs. 95%), overall results indicate its reliability. This suggests the system operates effectively without constant instructor oversight and is scalable for large or online classes.

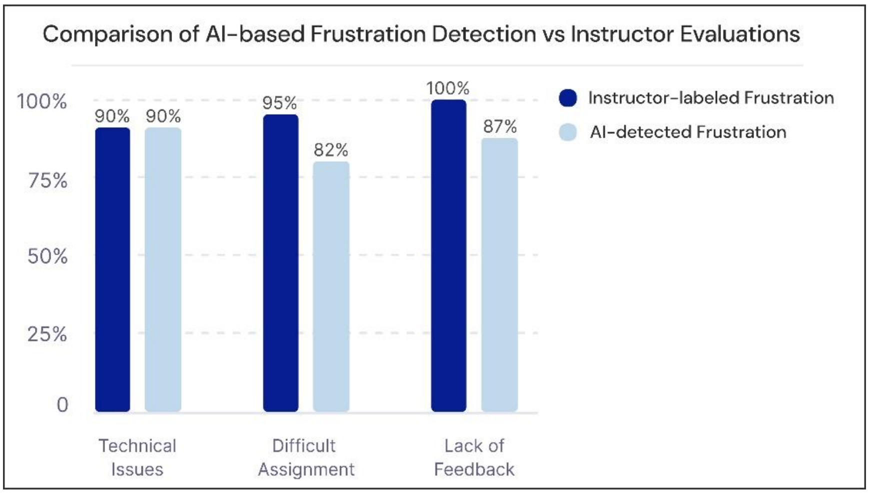

Figure 4: AI vs. Instructor Frustration Detection

4.2 Comparative Analysis of Model Variations

To evaluate the effectiveness of the multimodal approach, experiments were conducted on the same group of students, with each modality activated sequentially before combining them. Frustration detection was first assessed using facial expressions, then vocal cues, and finally both together in a multimodal fusion model. This sequential approach enabled a direct comparison, confirming the advantages of multimodal fusion. As shown in Table 5, the fusion model outperformed unimodal models by ~10% in accuracy, demonstrating that frustration is best detected through a combination of facial and vocal signals rather than a single modality alone.

Among single-modality models, the audio model (78%) slightly outperformed the facial model (75%), suggesting that vocal cues were more reliable for frustration detection. This could be due to facial masking in formal settings, while vocal tone is harder to control. Additionally, DAiSEE's frustration labels may have included milder cases that were visually ambiguous, whereas IEMOCAP's frustration-labeled utterances were clearer. The attention-based fusion further improved performance (F1 = 0.85) by dynamically adjusting modality weights, favoring voice when facial cues were neutral and vice versa. Multimodal fusion improved detection confidence and reliability. Notably, hidden frustration misses were lowest with the fusion model, indicating better capture of nuanced emotional signals.

Table 5: Analysis of Frustration Detection by Parameter During the Experimental Phase

	Performance				Errors	Confidence	
	Accuracy	Precision	Recall	F1-score	Hidden Frustration Misses	Average Confidence (Conf. Avg)	Max Confidence (Conf. Max)
Facial CNN (Vision)	77%	0.79	0.70	0.74	7	0.71	0.92
Audio Transformer	76%	0.79	0.76	0.77	6	0.74	0.95
Multimodal Fusion	84%	0.84	0.86	0.85	10	0.81	0.98
+% from CNN	**+9.1%**	**+6.3%**	**+22.9%**	**+14.9%**	**+42.9%**	**+14.1%**	**+6.5%**
+% from Audio	**+10.5%**	**+6.3%**	**+13.2%**	**+10.4%**	**+66.7%**	**+9.5%**	**+3.2%**

Error analysis identified key challenges in unimodal models. The facial model often misclassified concentration as frustration, while the audio model confused high-arousal emotions like excitement with frustration. The fusion model reduced these errors by leveraging cross-modal context, though misclassifications persisted when both modalities misaligned, such as strained vocal tones paired with a furrowed brow. Transformer-based audio processing outperformed LSTM (~2-3% higher F1-score), likely due to its ability to capture key vocal cues early. The model's decision window (2-5s) balanced reactivity and stability, avoiding the noise of shorter windows while preventing missed transient frustration signals in longer ones.

This analysis directly addresses the research questions (RQ).

RQ1: Can a multimodal AI model significantly improve frustration detection compared to single-modal approaches?

The findings demonstrate that the multimodal AI model (Accuracy = 85%, F1-score = 0.85) significantly outperforms unimodal approaches, including facial expression analysis alone (75%) and speech emotion recognition alone (78%). These results highlight the advantage of integrating both modalities, as they provide complementary information that enhances the reliability of frustration detection.

RQ2: How well do automated predictions align with human (instructor) assessments of frustration in real learning scenarios?

The study demonstrates a strong alignment between AI-based frustration detection and instructor evaluations, with a precision of 84% and recall of 86%. This suggests that the AI system is highly effective in identifying frustration when clear affective cues are present. However, some limitations persist, particularly in cases where frustration is internalized or subtly expressed, leading to occasional false negatives.

RQ3: What are the practical benefits and challenges of implementing such technology in educational settings?

The system's ability to detect frustration in real-time suggests its potential integration into adaptive learning platforms. However, effective implementation requires careful calibration of detection thresholds to balance sensitivity and specificity, minimizing false alarms while ensuring meaningful intervention opportunities. Additionally, ethical considerations related to privacy, consent, and potential bias must be addressed to facilitate responsible deployment in educational environments.

RQ4: How can machine-driven frustration detection assist educators in identifying students who may need additional psychological support or consultation?

The system's capacity to detect early indicators of frustration enables educators to identify students at risk of disengagement or heightened emotional distress. This proactive approach can support timely psychological interventions, fostering a more supportive learning environment. Future refinements, particularly in context-aware modeling and personalized sensitivity adjustments, will further enhance the system's capacity to assist educators in addressing students' emotional and academic needs.

4.3 Interpretation of Errors and Model Limitations

While the model effectively detected frustration, certain limitations led to errors. Key issues include: (1) *Confusion vs. Frustration Misclassification:* The model often mistook confusion for frustration, especially in students concentrating intensely. Expressions like furrowed brows and squinting appeared in both states, making differentiation difficult. The binary classification approach did not explicitly account for confusion, leading to misclassifications. A multi-label model or task performance data (e.g., tracking incorrect attempts) could help distinguish these emotions. (2) *Subtle or Internalized Frustration Misses:* Some students exhibited frustration in subtle ways, such as posture changes, which the AI struggled to detect. Unlike a human instructor who could infer frustration from context, the model lacked situational awareness. Temporal smoothing, while reducing false positives, sometimes ignored brief frustration moments. A more context-aware approach could improve detection of mild frustration. (3) Short-lived or Contextual Frustration: The model flagged temporary frustration that quickly resolved, such as brief reactions to technical issues. Since it did not track frustration duration, alerts were sometimes unnecessary. Future refinements could incorporate frustration persistence tracking (e.g., only triggering alerts if frustration lasts over 30 seconds). (4) Sensor/Input Limitations: Real-world conditions, such as face occlusion, poor lighting, or background noise, impacted detection. In group settings, isolating a single student's voice was challenging, leading to misattributions (e.g., detecting another student's sigh as frustration). The audio model also struggled with soft-spoken frustration, as it prioritized vocal intensity. Integrating speech-to-text analysis could help by identifying explicit frustration-related words rather than relying solely on tone.

These limitations are consistent with the Cognitive Disequilibrium Theory (Graesser and D'Mello, 2012), which suggests that frustration and confusion lie on a continuum, where unresolved confusion may develop into frustration. This may partly explain the model's difficulty in distinguishing between the two. A lack of contextual awareness also led to occasional false positives unrelated to coursework. Incorporating task-related metrics, such as incorrect attempts or delayed responses, and applying adaptive thresholds based on individual expressiveness could improve detection accuracy. Although these refinements were not implemented due to sample size constraints, the system still provides valuable early indicators to support instructors.

4.4 Ethical Considerations and Bias in AI-Based Emotion Detection

The study involved university students from Kazakhstan, representing both European and Central Asian ethnicities. This reflects local diversity and helped calibrate the model to common facial and vocal expressions in the region. While performance was consistent across tested groups, generalizability to other populations remains uncertain due to cultural and linguistic differences in emotional expression.

The study identified no major biases across gender or ethnicity, though the limited sample size prevents definitive conclusions. The model may still underperform for individuals expressing frustration atypically, including culturally diverse or neurodivergent students. This echoes concerns from AI proctoring systems about fairness and unintended impacts (Sakhipov, Omirzak and Fedenko, 2025). Both false positives and missed detections could affect student support, reinforcing the need for larger, more varied datasets and ongoing evaluation.

Participants sometimes altered their behavior due to awareness of being monitored, highlighting the importance of minimizing observer effects. Emotion detection tools should assist, not surveil, with final decisions left to educators. Ethical deployment requires full transparency, informed consent, and compliance with GDPR and UNESCO principles to ensure fairness, privacy, and student trust.

4.5 Implications and Applications

The promising performance of the frustration detection model opens up several applications in adaptive learning systems. One immediate use is real-time alerts for instructors. In in-person or online classroom settings, a dashboard could highlight students displaying frustration, such as by placing a red border around their video feed, allowing teachers to check in before a student disengages. This early warning system would help guide targeted interventions, enabling more proactive classroom management. An early version of the frustration analysis interface is shown in Figure 5; future versions may introduce design improvements and expand visualization capabilities.

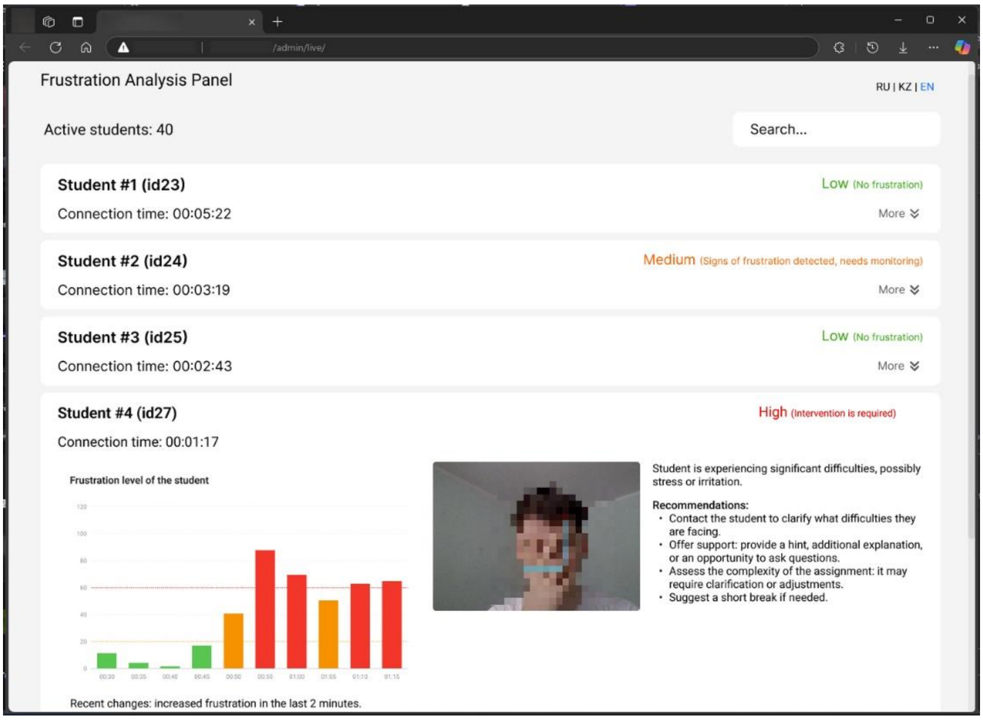

Figure 5: Real-Time Frustration Analysis Panel: Student Frustration Detection with Intervention Recommendations

The AI model can enhance adaptive tutoring by adjusting content based on frustration detection. When a student struggles, it can offer hints, encouragement, or modify the difficulty level to maintain learning flow. Self-awareness feedback could help students regulate emotions by suggesting breaks or review sessions. Aggregated frustration data informs curriculum improvements, highlighting lessons that need refinement. Tracking frustration trends enables personalized support, while emotion data can optimize group learning by balancing teams. Expanding detection to boredom or confusion could refine engagement strategies. Real-time frustration indicators and adaptive feedback foster student engagement, emotional regulation, and curriculum optimization.

The findings highlight the practical viability of multimodal AI systems for real-time frustration detection in educational environments. Technologically, the proposed model can be integrated into adaptive learning platforms, enabling timely and personalized interventions. Pedagogically, it supports instructors in identifying at-risk students, facilitating early emotional and academic support. At the institutional level, aggregated frustration trends may inform curriculum design and psychological service allocation. The findings provide a foundation for implementing frustration-aware feedback in intelligent tutoring systems and for guiding institutional strategies that enhance student support and well-being.

5. Conclusion

This study demonstrates the potential of AI-based emotion recognition to support education through a multimodal system that detects student frustration using facial and vocal cues. The model outperformed unimodal approaches and showed high accuracy in both controlled and real-world conditions, offering practical benefits for adaptive learning. By providing real-time emotional feedback, the system can help instructors identify at-risk students and respond with timely support, especially in digital or large-scale classrooms where individual monitoring is limited. In addition to individual interventions, such tools may also contribute to a more emotionally supportive classroom climate by making students' affective states more visible and actionable for educators.

The findings underscore the importance of integrating multiple modalities to capture both overt and subtle expressions of frustration. While the system shows promise, it was tested in structured scenarios and uses binary classification, which may oversimplify emotional dynamics. Its broader effectiveness and long-term impact on learning remain to be explored. Future work will also focus on refining multi-label classification, expanding cross-cultural validation, and embedding emotion-aware systems into diverse educational technologies. Educators and edtech developers may consider integrating such systems as part of responsive instructional design, guided by clear protocols for transparency, consent, and student support.

6. Limitations and Future Work

6.1 Delimitations of the Study

This study intentionally focused on the detection of frustration as a target emotional state, excluding other related emotions such as confusion, boredom, or anxiety. The participant pool was limited to university students in Kazakhstan to ensure contextual consistency and manageability, and findings may not generalize to other demographics or educational settings. Only facial and vocal modalities were used for emotion recognition; behavioral logs, physiological data, and contextual cues were deliberately excluded. The system was tested in short, individual learning sessions, and long-term emotional trends or academic outcomes were not examined. Additionally, considerations such as data storage, privacy, and large-scale deployment in real classroom environments were beyond the scope of this initial study, which was primarily focused on developing and validating the core AI model. These aspects are recognized as important directions for future research and practical implementation.

6.2 Research Limitations

The sample consisted of 160 university students from Kazakhstan, reflecting local ethnic diversity. While this ensures cultural relevance, it limits generalizability to other populations where frustration may be expressed differently. The model was also trained on benchmark datasets containing acted emotional responses, which may not fully reflect how frustration occurs in real educational settings. As a result, the model's behavior in natural classroom environments remains untested.

The system was evaluated using binary classification, which simplifies emotional states and does not capture transitions between related emotions such as confusion or disengagement. In addition, the study did not examine whether using frustration detection leads to improved academic performance, motivation, or persistence. This remains an open question for future applied research. Although behavioral context such as task progress was observed during annotation, the model itself does not use these signals as input. Incorporating them could improve detection of subtle or internalized frustration.

6.3 Future Directions

Future research should explore the impact of AI-based frustration detection on academic outcomes, including learning performance, task completion, and knowledge retention. Experimental studies could assess whether real-time emotional feedback enhances student engagement, motivation, or instructional responsiveness. It is also important to examine effects on metacognitive regulation, help-seeking behavior, and persistence in cognitively demanding contexts. Embedding the system into existing educational platforms would allow for usability testing and evaluation of how instructors and learners interpret and apply emotional insights. Further work should address generalizability by testing across age groups, languages, and educational systems, as well as through longitudinal designs that capture emotional dynamics over time. Improving detection sensitivity may require integrating behavioral indicators such as gaze, response timing, or interaction patterns. Models that account for emotional transitions could support more adaptive and context-aware feedback. Finally, the pedagogical use of emotion data must be guided by principles of transparency, privacy, and student autonomy, requiring close collaboration between AI developers, educators, and institutional stakeholders.

AI Statement: AI tools were not used for writing or generating the content of this research paper. Artificial intelligence was only employed during model development and performance evaluation for frustration detection. All research design, data analysis, and manuscript preparation were conducted manually by the authors.

Ethics Statement: This study followed strict privacy and confidentiality standards. No personally identifiable information was collected, stored, or accessed by the researchers at any point. Facial and vocal inputs were processed locally and automatically by the AI model in real time, with no human access, recording, or transmission. No raw biometric or behavioral data were retained or available during or after the study. All analysis was based on fully anonymized outputs. Participants provided informed consent prior to participation, and the study was designed to ensure minimal intrusion and no harm. The research complies with the GDPR, ensuring lawful, fair, and transparent processing aligned with privacy-by-design principles for AI applications in education.

References

Abbaschian, B.J., Sierra-Sosa, D. and Elmaghraby, A., 2021. Deep learning techniques for speech emotion recognition, from databases to models. *Sensors*, 21(4), p.1249. https://doi.org/10.3390/s21041249.

Aguilera, A., Mellado, D. and Rojas, F., 2023. An assessment of in-the-wild datasets for multimodal emotion recognition. *Sensors*, 23(11), p.5184. https://doi.org/10.3390/s23115184

Agung, E.S., Rifai, A.P. and Wijayanto, T., 2024. Image-based facial emotion recognition using convolutional neural network on emognition dataset. *Scientific Reports*, 14, p.14429. https://doi.org/10.1038/s41598-024-65276-x

Ahmed, N., Al Aghbari, Z. and Girija, S., 2023. A systematic survey on multimodal emotion recognition using learning algorithms. *Intelligent Systems with Applications*, 17, p.200171. https://doi.org/10.1016/j.iswa.2022.200171

Baker, R.S., Cloude, E., Andres, J.M.A.L. and Wei, Z., 2025. The confrustion constellation: A new way of looking at confusion and frustration. *Cognitive Science*, 49(1), e70035. https://doi.org/10.1111/cogs.70035.

Banzon, A.M., Beever, J. and Taub, M., 2024. Facial expression recognition in classrooms: Ethical considerations and proposed guidelines for affect detection in educational settings. *IEEE Transactions on Affective Computing*, 15(1), pp.93–104. https://doi.org/10.1109/TAFFC.2023.3275624.

Busso, C., Bulut, M., Lee, C.-C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J.N., Lee, S. and Narayanan, S.S., 2008. IEMOCAP: Interactive emotional dyadic motion capture database. *Language Resources and Evaluation*, 42(4), pp.335–359. https://doi.org/10.1007/s10579-008-9076-6

Bustos-López, M., Cruz-Ramírez, N., Guerra-Hernández, A., Sánchez-Morales, L.N., Cruz-Ramos, N.A. and Alor-Hernández, G., 2022. Wearables for engagement detection in learning environments: A review. *Biosensors*, 12(7), p.509. https://doi.org/10.3390/bios12070509.

Cheng, Z., Cheng, Z.-Q., He, J.-Y., Sun, J., Wang, K., Lin, Y., Lian, Z., Peng, X. and Hauptmann, A., 2024. Emotion-LLaMA: Multimodal emotion recognition and reasoning with instruction tuning. *arXiv preprint*. Available at: https://arxiv.org/abs/2406.11161 [Accessed 16 June 2025].

Corza-Vargas, V.M., Martinez-Maldonado, R., Escalante-Ramirez, B. and Olveres-Montiel, J., 2024. Students' ethical, privacy, design, and cultural perspectives on visualizing cognitive-affective states in online learning. *Journal of Learning Analytics*, 11(3), pp.24–40. https://doi.org/10.18608/jla.2024.8483.

Graesser, A.C. and D'Mello, S., 2012. Chapter five - Emotions during the learning of difficult material. In: B.H. Ross, ed. *Psychology of Learning and Motivation*. Vol. 57. Academic Press, pp.183-225. https://doi.org/10.1016/B978-0-12-394293-7.00005-4.

Gupta, A., D'Cunha, A., Awasthi, K. and Balasubramanian, V., 2022. DAiSEE: Towards user engagement recognition in the wild. *arXiv preprint*. Available at: https://arxiv.org/abs/1609.01885v7 [Accessed 10 March 2025].

Henderson, N., Min, W., Rowe, J. and Lester, J., 2021. Enhancing multimodal affect recognition with multi-task affective dynamics modeling. *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Nara, Japan, pp.1–8. https://doi.org/10.1109/ACII52823.2021.9597432.

Khare, S.K., Blanes-Vidal, V., Nadimi, E.S. and Acharya, U.R., 2024. Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. *Information Fusion*, 102, p.102019. https://doi.org/10.1016/j.inffus.2023.102019

Lian, H., Lu, C., Li, S., Zhao, Y., Tang, C. and Zong, Y., 2023. A survey of deep learning-based multimodal emotion recognition: Speech, text, and face. *Entropy*, 25(10), p.1440. https://doi.org/10.3390/e25101440

Liu, Z., Xie, Y., Sun, Z., Liu, D., Yin, H., and Shi, L., 2023. Factors associated with academic burnout and its prevalence among university students: a cross-sectional study. *BMC Med Educ,* 23, 317. https://doi.org/10.1186/s12909-023-04316-y.

Livingstone, S.R. and Russo, F.A., 2018. The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PLOS ONE*, 13(5), e0196391. https://doi.org/10.1371/journal.pone.0196391.

Malekshahi, S., Kheyridoost, J.M. and Fatemi, O., 2024. A general model for detecting learner engagement: Implementation and evaluation. *arXiv preprint*. Available at: https://doi.org/10.48550/arXiv.2405.04251 [Accessed 10 March 2025].

Mamieva, D., Abdusalomov, A.B., Kutlimuratov, A., Muminov, B. and Whangbo, T.K., 2023. Multimodal emotion detection via attention-based fusion of extracted facial and speech features. *Sensors (Basel, Switzerland)*, 23(12), p.5475. https://doi.org/10.3390/s23125475.

Mattioli, M. and Cabitza, F., 2024. Not in my face: Challenges and ethical considerations in automatic face emotion recognition technology. *Machine Learning and Knowledge Extraction*, 6(4), pp.2201-2231. https://doi.org/10.3390/make6040109.

Moon, J., Ke, F., Sokolikj, Z. and Dahlstrom-Hakki, I., 2022. Multimodal data fusion to track students' distress during educational gameplay. *Journal of Learning Analytics*, 9(3), pp.75-87. https://doi.org/10.18608/jla.2022.7631.

Novak, E., McDaniel, K. and Li, J., 2023. Factors that impact student frustration in digital learning environments. *Computers and Education Open*, 5, p.100153. https://doi.org/10.1016/j.caeo.2023.100153.

Pekrun, R. and Marsh, H.W., 2022. Research on situated motivation and emotion: Progress and open problems. *Learning and Instruction*, 81, p.101664. https://doi.org/10.1016/j.learninstruc.2022.101664.

Pham, T.-D., Duong, M.-T., Ho, Q.-T., Lee, S. and Hong, M.-C., 2023. CNN-based facial expression recognition with simultaneous consideration of inter-class and intra-class variations. *Sensors*, 23(24), p.9658. https://doi.org/10.3390/s23249658

Pordoy, J., Farman, H., Dicheva, N.K., Anwar, A., Nasralla, M.M., Khilji, N. and Rehman, I.U., 2024. Multi-frame transfer learning framework for facial emotion recognition in e-learning contexts. *IEEE Access*, 12, pp.151360–151381. https://doi.org/10.1109/ACCESS.2024.3478072.

Qian, F. and Han, J., 2022. Contrastive regularization for multimodal emotion recognition using audio and text. *arXiv preprint* arXiv:2211.10885. Available at: https://arxiv.org/abs/2211.10885 [Accessed 16 June 2025].

Rahman, M.M., Munir, M.U., Rahman, M.M. and Badiuzzaman, M., 2024. EmoDetect: A learning-centred affective database for detecting student frustration in online learning. *2024 5th International Conference on Advancements in Computational Sciences (ICACS)*. Lahore, Pakistan, pp.1-6. https://doi.org/10.1109/ICACS60934.2024.10473235.

Rhue, L., 2018. Racial influence on automated perceptions of emotions. *SSRN*. Available at: https://ssrn.com/abstract=3281765 [Accessed 4 March 2025].

Sakhipov, A., Yermaganbetova, M., Latypov, R. and Ualiyev, N., 2022. Application of blockchain technology in higher education institutions. *Journal of Theoretical and Applied Information Technology*, 100(4), pp.1138–1147.

Sakhipov, A., Omirzak, I. and Fedenko, A., 2025. Beyond face recognition: A multi-layered approach to academic integrity in online exams, *Electronic Journal of e-Learning*, 23(1), pp.81-95. https://doi.org/10.34190/ejel.23.1.3896.

Shou, Z., Huang, Y., Li, D., Feng, C., Zhang, H., Lin, Y. and Wu, G., 2024. A student facial expression recognition model based on multi-scale and deep fine-grained feature attention enhancement. *Sensors*, 24(20), p.6748. https://doi.org/10.3390/s24206748.

Solanki, N. and Mandal, S., 2022. Engagement analysis using DAiSEE dataset. *2022 17th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Singapore, pp.223-228. https://doi.org/10.1109/ICARCV57592.2022.10004250.

Song, M., Parada-Cabaleiro, E., Liu, S., Milling, M., Baird, A., Yang, Z. and Schuller, B.W., 2021. Supervised contrastive learning for game-play frustration detection from speech. *In: M. Antona and C. Stephanidis, eds. Universal Access in Human-Computer Interaction. Design Methods and User Experience*. Cham: Springer International Publishing, pp.617–629. https://doi.org/10.1007/978-3-030-78092-0_43.

Villegas-Ch, W.E., García-Ortiz, J. and Sánchez-Viteri, S., 2023. Identification of emotions from facial gestures in a teaching environment with the use of machine learning techniques. *IEEE Access*, 11, pp.38010–38022. https://doi.org/10.1109/ACCESS.2023.3267007.

Wang, C., 2022. Emotion recognition of college students' online learning engagement based on deep learning. *International Journal of Emerging Technologies in Learning (iJET)*, 17(06), pp.110–122. https://doi.org/10.3991/ijet.v17i06.30019.

Wang, Y., Gu, Y., Yin, Y., Han, Y., Zhang, H., Wang, S., Li, C. and Quan, D., 2023. Multimodal transformer augmented fusion for speech emotion recognition. *Frontiers in Neurorobotics*, 17, 1181598. https://doi.org/10.3389/fnbot.2023.1181598.

Xiang, A., Qi, Z., Wang, H., Yang, Q. and Ma, D., 2024. A multimodal fusion network for student emotion recognition based on transformer and tensor product. *arXiv preprint*, arXiv:2403.08511. Available at: https://arxiv.org/abs/2403.08511 [Accessed 16 June 2025].

Yan, J., Li, P., Du, C., Zhu, K., Zhou, X., Liu, Y. and Wei, J., 2024. Multimodal emotion recognition based on facial expressions, speech, and body gestures. *Electronics*, 13(18), p.3756. https://doi.org/10.3390/electronics13183756

Zaidi, S.A.M., Latif, S. and Qadir, J., 2023. Cross-language speech emotion recognition using multimodal dual attention transformers. *arXiv preprint*. Available at: https://doi.org/10.48550/arXiv.2306.13804 [Accessed 10 March 2025].

Zaidi, S.A.M., Latif, S. and Qadir, J., 2024. Enhancing cross-language multimodal emotion recognition with dual attention transformers. *IEEE Open Journal of the Computer Society*, 5, pp.684–693. https://doi.org/10.1109/OJCS.2024.3486904