INNOEDUCA

International Journal of Technology and Educational Innovation

INNOEDUCA-GRUPO DE INVESTIGACIÓN
ISSN-e 2444-2925
Vol. 11 nº 1

umaeditorial

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - ISSN-e: 2444-2925

INNOEDUCA

PRESENTACIÓN

Innoeduca. International Journal of Technology and Educational Innovation es una publicación científica que nace auspiciada por el Grupo de investigación Innoeduca (grupo consolidado de la Junta de Andalucía - SEJ-533) de la Universidad de Málaga (España). Innoeduca es un grupo interdisciplinar de docentes e investigadores (pedagogos, matemáticos, informáticos, diseñadores gráficos…) de distintos niveles educativos, que desarrollan productos, investigaciones y formación en el campo de la Innovación y la Tecnología Educativa. Desde sus inicios, el grupo ha desarrollado una labor investigadora permanente y ha tenido como prioridades el contacto y la colaboración con otros investigadores y centros nacionales e internacionales.

Innoeduca. International Journal of Technology and Educational Innovation es una publicación en línea, abierta y revisada por pares, que proporciona una plataforma para exponer y compartir conocimientos en forma de artículos de investigación empírica y teórica, estudios de caso y revisión de la literatura. Los artículos enviados deberán ajustarse a las normas de publicación y tratar sobre educación, innovación y tecnología.

Esta publicación surge con un compromiso de rigor en el proceso editorial (selección de manuscritos, plazos de edición y calidad del resultado final) avalado por un comité científico de máximo prestigio internacional.

Difundir contenidos de calidad entre la comunidad científica es la finalidad de este proyecto. Por ello, se admitirán artículos escritos en inglés, español o portugués.

Esperamos que este número resulte interés al lector dada la relevancia de las investigaciones publicadas.

Julio Ruiz-Palmero
Director de Innoeduca. International Journal of Technology and Educational Innovation

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - ISSN-e: 2444-2925

ÍNDICE

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 5-28 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20219

Understanding ChatGPT Adoption among Higher Education Students in Punjab, India: An Application of UTAUT2 Model

Comprender la adopción de ChatGPT entre estudiantes de educación superior en Punjab, India: una aplicación del modelo UTAUT2

RECEIVED 02/07/2024 ACCEPTED 29/11/2024 PUBLISHED 01/06/2025

Sukhjit Singh
ORCID 0009-0008-6915-3423
Chitkara University, India
sukhjit.s@chitkara.edu.in

Pooja Singh
ORCID 0009-0004-6355-3208
Chitkara University, India
pooja.singh@chitkara.edu.in

Vismaad Kaur
ORCID 0000-0002-0112-4502
Chitkara University, India
vismaad.kaur@chitkara.edu.in

ABSTRACT

This study examines Indian Higher Education students' behavioural intention to use ChatGPT in their learning. Unified Theory of Acceptance, and Use of Technology 2 (UTAUT2) model is used to investigate the impact of the eight UTAUT2 factors on the students' behavioural intention towards using ChatGPT. A pilot study on 100 students was done to check the reliability and validity of the instrument based on the UTAUT2 model. Using a quantitative research approach, data was gathered from 362 students of Punjab (A North region State), India (313 students' data was included in final analysis) using purposive sampling technique. The study's findings revealed that PE (Performance Expectancy), SI (Social Influence), HM (Hedonic Motivation), Hb (Habit), FC (Facilitating Conditions) had significant positive influence on BI (Behavioural Intention) whereas EE (Effort Expectancy) had not significantly influenced BI. On ChatGPT use, H and BI had a positive influence, but FC did not significantly influence ChatGPT use. 67% of the respondents gave priority to learning AI tools in school. In terms of practical implications, this study adds to the current literature on ChatGPT or AI tools in higher education, being useful to education scholars. Also, this study highlights the validation of UTAUT2 model to use ChatGPT among HEI students in Punjab, India. The findings of this study could facilitate discussions among educators working for policies related to the use of AI tools, specifically ChatGPT in India.

KEYWORDS ChatGPT; Acceptance and Use; UTAUT2; PLS-SEM; Higher Education.

RESUMEN

Este estudio examina la intención conductual de los estudiantes indios de educación superior al utilizar ChatGPT en su aprendizaje. El modelo de Teoría Unificada de Aceptación y Uso de Tecnología 2 (UTAUT2) se utiliza para investigar el impacto de los ocho factores UTAUT2 en la intención de comportamiento de los estudiantes hacia el uso de ChatGPT. Se realizó un estudio piloto con 100 estudiantes para comprobar la confiabilidad y validez del instrumento basado en el modelo UTAUT2. Utilizando un enfoque de investigación cuantitativa, se recopilaron datos de 362 estudiantes de Punjab (un estado de la región norte), India (los datos de 313 estudiantes se incluyeron en el análisis final) utilizando una técnica de muestreo intencional. Los hallazgos del estudio revelaron que PE (expectativa de desempeño), SI (influencia social), HM (motivación hedónica), Hb (hábito), FC (condiciones facilitadoras) tuvieron una influencia positiva significativa en BI (intención conductual), mientras que EE (expectativa de esfuerzo) tuvo no influyó significativamente en el BI. En el uso de ChatGPT, H y BI tuvieron una influencia positiva, pero FC no influyó significativamente en el uso de ChatGPT. El 67% de los encuestados dio prioridad al aprendizaje de herramientas de IA en la escuela. En términos de implicaciones prácticas, este estudio se suma a la literatura actual sobre ChatGPT o herramientas de IA en la educación superior, siendo útil para los académicos de la educación. Además, este estudio destaca la validación del modelo UTAUT2 para utilizar ChatGPT entre estudiantes de IES en Punjab, India. Los hallazgos de este estudio podrían facilitar los debates entre los educadores que trabajan en políticas relacionadas con el uso de herramientas de inteligencia artificial, específicamente ChatGPT en India.

PALABRAS CLAVE ChatGPT; Aceptación y Uso; UTAUT2; PLS-SEM; Educación Superior.

1. INTRODUCTION

Artificial intelligence (AI) has entered almost each and every professional field in today's world and can be leveraged in the education field as well. It has opened a significant plethora of innovative opportunities for the teaching-learning processes and practices. From personalized learning experiences to predictive analytics, AI has had a profound impact on the way students learn and educators teach (Huang, 2023). With the advent of AI, and the increasing interest in the available technical AI chatbot applications, including Chat-GPT, Google Bard, Microsoft Bing, Jasper, and others, the students have been trying their hand at using these tools for finishing academic tasks at a faster pace. Everyone today needs quick fixes, and the students are always looking for ways to make their academic life easier and reduce related stress. Since ChatGPT's launch, generative artificial intelligence (gen-AI) systems have received a lot of attention because of its potential influence in a variety of domains (Dowling & Lucey, 2023; Eke, 2023; Lim et al., 2023; Soni et al., 2022; Vaishya et al., 2023), including higher education (Choi et al., 2023; Duong et al., 2023). ChatGPT has been downloaded more than a million times in just one week after its 2022 launch (Lund & Wang, 2023; Pavlik, 2023). ChatGPT has an enormous potential to enhance the effectiveness of learning activities, including creating customized content, assisting with assignments, and giving students feedback (Lund & Wang, 2023). But it also has a certain disadvantage as it could make the students totally dependent on it.

In the present era, students, Generation Z, "zoomers", or "digital natives" (Lim et al., 2022), are inclined towards integrating new technologies into their daily study routine. And this integration is not just limited to the higher education students. Students at all levels-primary, secondary, tertiary, and higher education, have begun using ChatGPT (Duong et al., 2023). ChatGPT has been used a lot in higher education (24.18%), K-12 education (22.09%), and practical skills learning (15.28%) as per Mogavi et al. (2024). This suggests that students

are using ChatGPT to reduce cognitive load, further affecting their creativity and thinking skills negatively. A.I. tools in 21st century are not only marking their presence in classroom learning but also AlKursheh (2024) also highlighted the importance of Artificial intelligence in assesssment specifically in Higher education.

Following a review of the existing literature, the analysis pinpoints six crucial elements that many writers have suggested that encourage readers to use text generative AIs such as ChatGPT. The identified factors are: Time Saving and Task Management (TSTM), Ease of Access (EA), Aided Learning (AL), Inseparability of Content (IC), Technical Knowledge of the Program (TKP), Cognitive Miserliness of the User (CMU) (Niloy et al. 2024). These variables point towards the fact that if students continue to rely on ChatGPT, the goal of helping them build 21st century or life skills may appear very distant. It is recommended that students utilise ChatGPT as an additional tool to enhance their research and learning. It should not be used in place of crucial learning components, such as engaging with primary sources, gaining varied viewpoints, and developing critical thinking abilities. Preventing over-reliance on AI would give students the room they need to build the strong cognitive abilities required for independent learning (Mogavi et al. 2024). Aldosari (2020) concludes that experts anticipate an optimistic future scenario where AI enhances various academic facets including education quality, guidance, assessment, and student support. Experts advocate for faculty readiness through training programs to maximize AI benefits.

There is currently less peer-reviewed research that focuses only on how students in Higher Education use ChatGPT, as noted by Strzelecki and ElArabawy (2023). However, it has generated a great deal of interest from a variety of stakeholder groups, including college students who utilize this AI-powered application to help them finish their assignments (Strzelecki & ElArabawy, 2023). The UTAUT is a significant theory for examining people's intentions regarding the acceptance and use of technology (Venkatesh et al., 2003). Due to a lack of studies in the context of Indian higher education, the present study focused on understanding ChatGPT adoption among higher education students in Punjab, India, and applied the UTAUT2 model in the context of use of ChatGPT in the students.

2. LITERATURE REVIEW

2.1. Unified theory of acceptance and use of technology (UTAUT2 Model)

Four key components served as the primary predictors of information technology usage intention and acceptability in the previous UTAUT1, which was created by Venkatesh et al. in 2003 (Narayan & Naidu, 2024; Zacharis & Nikolopoulou, 2022). These consisted of social influence, enabling conditions, performance expectancy, and effort expectancy. The UTAUT2 model was further expanded upon by Venkatesh et al. (2012) by incorporating three more constructs: hedonic motivation, price value, and habit. As mentioned by Narayan and Naidu (2024), Researchers who have studied technology acceptance and use (such as e-learning) in higher education contexts include Al-Fraihat et al. (2020), Farooq et al. (2017), and Kumar and Bervell (2019). These researchers have used the UTAUT1 model.

The UTAUT2 model, which was created later, has also been investigated in other situations, such as online gaming, artificial intelligence, internet/mobile banking, online gaming, online gaming, artificial intelligence and VR technology (Gansser & Reich, 2021; Kwateng et al., 2018). The theory's creators urged

researchers to investigate and validate their idea using a range of technologies, settings, and subjects (Narayan & Naidu, 2024; Zacharis & Nikolopoulou, 2022). By applying the UTAUT2 model in the context of higher education, educators and administrators can make informed decisions regarding the implementation of new technologies, design effective training programs, and create strategies to promote technology acceptance and use among both teachers and students (Chang et al., 2023). Talan et al. (2024) investigated the use and adoption of mobile learning tools among university students by using UTAUT2 model in higher education and highlighted the new ways and tools of learning among 21st century learners.

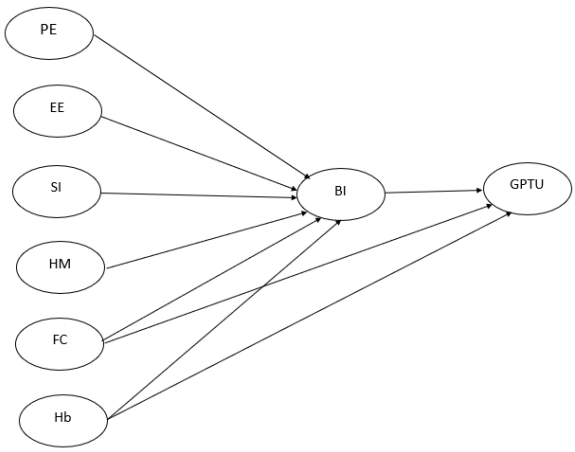
The UTAUT2 model was employed in a study on students at Indonesian higher education institutions to investigate the factors influencing the adoption and utilization of ChatGPT in the context of education. Facilitating situations were revealed to be the strongest predictor of behavioral intention to use ChatGPT in learning, according to the study. This implies that offering ChatGPT users the infrastructure, support, and resources they need can have a big impact on how ready students are to accept and use the technology. Additionally, the research revealed that the most significant factor influencing ChatGPT use was behavioral intention, suggesting that students' intentions to utilize the technology were critical to their actual utilization (Habibi et al., 2023).

2.2. Performance Expectancy (PE)

According to the UTAUT model, Performance Expectancy (PE) has a significant impact on an individual's inclination to use new technologies such as AI tools and ChatGPT for learning. According to Venkatesh et al. (2012), PE determines the user's ideas about their ability/capacity to use a specific technology (in this case, ChatGPT) to its full potential. Many studies in the context of education, specifically the use of AI technologies, indicated that PE was a strong predictor of BI (Andrews et al., 2021; Camilleri, 2024: Chatterjee & Bhattacharjee, 2020; Habibi et al., 2023; Lin et al., 2022). As Narayan and Naidu (2024) pointed out, some studies (Abbas, 2018; Ali et al., 2018; Althunibat, 2015; Dečman, 2015; Raza et al., 2021) found a positive association between PE and students' BI when adopting LMS, while other studies (Habibi et al., 2023; Zwain, 2019) found no positive impact.

H1: The significant role of PE in predicting BI to use CHATGPT among Higher Education Students of Punjab, India.

FIGURE 1. A Proposed Model Investigating Indian Hei Students' Chatgpt acceptance and use by UTAUT2 model



(Performance Expectancy **[PE]**, Effort Expectancy **[EE]**, Social Influence **[SI]**, Facilitating Conditions **[FC]**, Hedonic Motivation **[HM]**, Habit **[Hb]**, Behavioral Intention **[BI]**, ChatGPT Use **[GPTU]**)

In figure 1, list of hypotheses is : H1 – PE to BI, H2 – EE to BI, H3- SI to BI, H4- HM to BI, H5- FC to BI, H6- FC to GPTU, H7- Hb to BI, H8 – H to GPTU, H9- BI to GPTU.

2.3. Effort Expectancy (EE)

According to Venkatesh et al. (2012), the first UTAUT version defined EE as a variable that assesses how user-friendly technology-based tools and systems are. Positive relationships between EE and BI to apply AI have been observed in previous studies on a variety of technical systems, including digital libraries (Habibi et al., 2022), e-commerce (Bozorgkhou, 2015), banking systems (Abu-Taieh et al., 2022), and mobile payment (Al-Saedi et al., 2020). Education and artificial intelligence studies (Alhwaiti, 2023; Andrews et al., 2021; Chatterjee & Bhattacharjee, 2020; Guggemos et al., 2020; Lin et al., 2022; Raffaghelli et al., 2022) have looked into how EE affects BI. According to Guggemos et al. (2020) and Chatterjee & Bhattacharjee (2020), EE persuaded BI to use humanoid robots for academic writing and AI in Indian higher education, respectively. This PLS-SEM analysis demonstrates this influence.

H2: The significant role of EE in predicting BI to use CHATGPT among Higher Education Students of Punjab, India.

2.4. Social Influence (SI)

Social Influence (SI) refers to the extent to which influential individuals influence a person's decision to adopt technology (Venkatesh et al., 2003, 2012). Several research have indicated a favourable association between SI and students' intentions to utilize LMS (Im et al., 2011; Raza et al., 2021; Venkatesh et al., 2003) and video-based learning media assistance (Wijaya et al., 2022). As mentioned by Habibi et al. (2023), within the context of AI applications, including robotics (Guggemos et al., 2020), AI-based early warning systems (Raffaghelli et al., 2022), language learning with AI (Lin et al., 2022), augmented reality (Marto et al., 2019), post-COVID AI technology (Alhwaiti, 2023), and chatbots (Ragheb et al., 2022), researchers have reported that SI had a significant impact on BI.

H3: The significant role of SI in predicting BI to use CHATGPT among Higher Education Students of Punjab.

2.5. Hedonic Motivation (HM)

In the context of the UTAUT2 model, hedonic motivation is proposed as one of the constructs that influences behavioural intention and technology use. The UTAUT2 model extends the original UTAUT model by incorporating hedonic motivation, along with price value and habit, as additional constructs (Narayan & Naidu, 2024; Thongsri et al., 2018). Hedonic motivation refers to the pleasure or enjoyment that individuals derive from using a particular technology (Tamilmani et al., 2019). For example, as mentioned by Narayan & Naidu (2024), research like Raza et al. (2021) showed an insignificant impact of HM on BI in circumstances where in technology aids learning, Zwain (2019) and Hoi (2020) observed HM impacting both BI and Actual use. But Raza et al. (2020b) recommended that HM continue to be evaluated and examined for its impact on students' performance in further research. Study conducted by Habibi et al. (2023) concluded that Hm has a significant influence on BI. The integration of hedonic motivation into the UTAUT2 model provides

a comprehensive understanding of the factors that influence technology adoption and use in education, especially in the case of ChatGPT.

H4: The significant role of SI in predicting BI to use CHATGPT among Higher Education Students of Punjab, India.

2.6. Facilitating Conditions (FC)

Facilitating conditions (FC) is defined as the existence of suitable organizational infrastructure and resources to facilitate the deployment of technology (Narayan & Naidu, 2024; Zacharis & Nikolopoulou, 2022). FC in an educational setting involves students' access to technology gadgets such as Wi-Fi, high-speed internet broadband service, personal computers, smartphones, technical support, laboratories with the necessary equipment. Consequently, FC in this way, raises the BI and Use of e-learning technologies, improving student performance and e-learning system acceptance. An obvious argument which comes out of this situation that without such prompt resource support, educators and students will get demotivated. (Yeop et al., 2016). As noted by Habibi et al. (2023) and Narayan and Naidu (2024), the literature reveals considerable discrepancies between earlier research on the effect of FC on BI, just like it does for PE and EE. This could be a result of the technological, economic, and developing states of many nations and academic institutions.

According to Habibi et al. (2023), FC has a considerable impact on BI and GPTU among Indonesian higher education students. According to Binyamin and Zafar's (2021) systematic review, FC and BI are strongly linked in the context of Mobile Health Services. On the contrary, Raza et al. (2021) and Venkatesh et al. (2003) discovered that FC is insignificant.

H5: The significant role of FC in predicting BI to use CHATGPT among Higher Education Students of Punjab, India.
H6: The significant role of FC in predicting GPTU use CHATGPT among Higher Education Students of Punjab, India.

2.7. Habit (Hb)

Habit is one significant variable that was included in the UTAUT2 model (Gansser & Reich, 2021). As rightly mentioned by Habibi et al. (2023), the sustainable use of a technology-based system is habitual, meaning it is done without conscious thought or purpose. The automatic, recurring behavioural patterns that people acquire over time are referred to as habits. When it comes to AI tools like ChatGPT and other technologies, user adoption and use are greatly influenced by habit. Studies have indicated that the presence of a habit can have a substantial impact on a person's inclination to utilize technology (Habibi et al.,2023; Moorthy et al., 2018; Shivdas et al., 2020; Tseng et al., 2019; Venkatesh et al., 2012). Due to their accustomed ways of engaging with technology, people often rely on habits when utilizing AI technologies (Venkatesh et al., 2002).

Two hypotheses in relation to habit as also checked by Habibi et al. (2023).

H7: The significant role of Hb in predicting BI to use CHATGPT among Higher Education Students of Punjab, India.
H8: The significant role of Hb in predicting GPTU (Use of ChatGPT) among Higher Education Students of Punjab, India.

2.8. Behavioural Intention (BI)

Behavioural intention refers to an individual's preparedness to use a specific technology for various tasks (Venkatesh et al., 2003). Behaviour Intention is a measure of attitudes, beliefs, and behaviours (Salifu et al., 2024). Furthermore, beliefs about an innovation before to adoption can act as the most reliable predictor of intent to succeed or fail (Vishwanath & Goldhaber, 2003).

As highlighted by Salifu et al. (2024), based on previous literature, Attitudes and beliefs are commonly recognized as the driving forces behind a user's behavioral intention to do a specific activity, which may eventually lead to the use of a specific technology. On the other hand, Actual use behavior refers to the physical evidence of technology integration into an individual's everyday life, and it serves as a measure of the effectiveness of intentions in the real application of the technology (Habibi et al., 2023; Narayan & Naidu, 2023; Salifu et al., 2024). Based on the previous literature (Ajzen, 1991; Mustafa et al., 2022; Venkatesh et al., 2012), an individual's actual use of a new technology is frequently reflective of their behavioural intentions. Therefore, behavioural intention is the most important aspect in real usage. Also, previous literature (Motaghian et al., 2013; Habibi et al., 2023; Raza et al., 2020a, 2022; Wang & Wang, 2009) highlights the strong relationship between BI and actual use in technology-assisted learning.

H9: The significant role of BI in predicting GPTU (use of CHATGPT) among Higher Education Students of Punjab, India.

3. METHOD

3.1. Instrument used

The current study took into account the 31 items utilized by Habibi et al. (2023) on a 5-point Likert scale. Four items for PE, EE, SI, FC, and BI (Habibi et al., 2023; Venkatesh et al., 2012). Meanwhile, HM and GPTU comprise three items, while H contains five items (Habibi et al., 2023). Ten specialists in educational technology were invited to analyse the items for the content validity process. Finally, 31 items presented in Appendix 1 were used for both the pilot research and the full data collection.

3.2. Data Collection

Data collection for this study was done by non-probability sampling using a Google Form Survey. The questionnaire included sections of demographic information and the main survey questions. No personal information was recorded so that respondents could freely express their options without any fear of issues with confidentiality. In a pilot study to check the reliability of the adapted questionnaire, 100 responses were taken. The Cronbach alpha value of reliability of each variable came out to be > 0.7 (Ramu et al., 2023). These responses were not included in the final analysis. For final data collection, out of 600 google forms reached to respondents via contacting concerned Faculty members of the institutions in Punjab State, India. 362 responses were received for further analysis. Out of 362, 49 responses were discarded due to lack of heterogeneity in the responses. The standard deviation for these responses came out to be 0, so for the final analysis,

313 responses were considered. No personal details were recorded beyond gender and educational level (graduate or postgraduate) of the students to make responses authentic and to ensure the confidentiality of respondents. Among 313 respondents, in terms of Gender, 194 were female. (62%) and 119 males (38%) and, in terms of education level, 103 graduate students (33%) and 210 postgraduate students (67%).

3.3. Data analysis

To do the data analysis, PLS-SEM by using Smart PLS 4 was employed due to its robustness of handling sample sizes and non – normal data assumptions (Hair et al., 2021). Measurement and Structural Model were assessed along with IPMA (Importance-performance map analysis technique) to support the Structural Model. Moreover, IPMA is helpful to understand results of PLS-SEM by taking the performance of each construct into account (Ringle et al., 2018).

3.4. Data preparation

Table 1 shows the descriptive level information of items under variables i.e. PE, EE, SI, FC, HM, Hb, BI, GPTU which shows skewness and kurtosis values. Expect items Hb3 and Hb4 under the variable of Habit, all items have VIF less than 4 (Ramayah et al., 2018).

TABLE 1. Mean, Standard Deviation, Kurtosis, Skewness and VIF (Item wise)

Variable	Item	Mean	SD	Kurtosis	Skewness	VIF
PE	PE1	4.026	0.719	-0.297	-0.297	1.978
	PE2	3.965	0.751	0.383	-0.488	2.181
	PE3	3.933	0.771	0.766	-0.602	2.223
	PE4	3.923	0.784	0.626	-0.585	2.107
EE	EE1	4.042	0.734	0.586	-0.553	1.689
	EE2	3.802	0.766	0.662	-0.545	1.716
	EE3	4.051	0.701	1.344	-0.689	1.929
	EE4	3.706	0.863	0.642	-0.623	1.483
SI	SI1	3.706	0.821	0.063	-0.383	2.428
	SI2	3.581	0.854	0.031	-0.225	2.413
	SI3	3.396	0.906	0.106	-0.245	1.856
	SI4	3.307	0.952	-0.125	-0.266	1.887
FC	FC1	4.105	0.794	1.679	-0.998	1.713
	FC2	3.843	0.845	0.343	-0.559	1.591
	FC3	4.019	0.741	0.287	-0.457	1.728
	FC4	3.623	0.867	0.044	-0.368	1.315
HM	HM1	3.552	0.911	0.193	-0.416	2.704
	HM2	3.613	0.876	0.618	-0.565	3.567
	HM3	3.461	0.886	0.347	-0.349	2.814

Variable	Item	Mean	SD	Kurtosis	Skewness	VIF
Hb	Hb1	2.895	1.107	-0.695	0.125	3.041
	Hb2	3.083	1.075	-0.599	0.066	3.729
	Hb3	3.096	1.047	-0.477	-0.042	*5.394*
	Hb4	3.019	1.057	-0.565	0.027	*4.933*
	Hb5	3.278	1.051	-0.412	-0.309	2.105
BI	BI1	3.381	0.918	0.078	-0.278	1.856
	BI2	3.521	0.865	0.357	-0.392	2.665
	BI3	3.559	0.825	0.194	-0.224	3.168
	BI4	3.612	0.804	0.355	-0.214	2.817
GPTU	GPTU1	3.153	1.018	-0.563	-0.074	1.809
	GPTU2	2.728	1.099	-0.567	0.235	2.194
	GPTU3	2.99	1.047	-0.477	0.086	1.965

After deleting Hb3 and Hb4, new VIF for Hb1, Hb2 and Hb5 was 2.542, 2.576 and 1.857, respectively.

4. FINDINGS

4.1. Measurement Model

The measurement model was employed to assess the reliability and validity of the proposed model (Figure 1). As mentioned by Hair et al. (2019), Loadings above 0.708 are recommended. Table 2 shows that loadings are above this threshold. Values of Cronbach's alpha, rho_A and Composite Reliability for variables in the model are within threshold limits (higher than 0.700). In table 2, the measure of convergent validity i.e. AVE (Average Variance Explained) is higher than >50% of the variance which indicates that each construct in the model explains 50% or more of the variance of the items that make up the construct (Hair et al., 2019). HM has highest AVE with 83% of the variance explained with FC has lowest value with 60% variance explained.

TABLE 2. Loadings, Cronbach alpha, rho_A, CR, AVE of Variables

Variable	Item	Loadings	Alpha	rho_A	CR	AVE
BI	BI1	0.811	0.892	0.894	0.926	0.757
	BI2	0.886				
	BI3	0.903				
	BI4	0.878				
EE	EE1	0.742	0.808	0.828	0.872	0.631
	EE2	0.788				
	EE3	0.85				
	EE4	0.794				
FC	FC1	0.802	0.783	0.782	0.86	0.607
	FC2	0.79				
	FC3	0.801				
	FC4	0.719				

Variable	Item	Loadings	Alpha	rho_A	CR	AVE
GPTU	GPTU1	0.869	0.836	0.844	0.901	0.752
	GPTU2	0.874				
	GPTU3	0.859				
Hb	Hb1	0.901	0.862	0.862	0.916	0.784
	Hb2	0.902				
	Hb5	0.853				
HM	HM1	0.898	0.904	0.908	0.94	0.839
	HM2	0.936				
	HM3	0.914				
PE	PE1	0.817	0.867	0.871	0.909	0.715
	PE2	0.858				
	PE3	0.857				
	PE4	0.852				
SI	SI1	0.825	0.829	0.835	0.886	0.661
	SI2	0.829				
	SI3	0.781				
	SI4	0.815				

For Discriminant Validity, as mentioned by Hair et al. (2019), according to Henseler et al. (2014), HTMT values for each construct need to be less than .90 to prove the constructs to be conceptually different. Table 3 shows that HTMT Values of all variables are within limits. Loadings in Table 5 for each variable are greater than their cross loadings on other variables. Loadings are highlighted in bold in Table 4. As per Fornell – Larcker criteria (Fornell & Larcker, 1981), the shared variance for all constructs under the model need to have smaller value than their AVEs (mentioned in Table 5).

TABLE 3. HTMT

	BI	EE	FC	GPTU	Hb	HM	PE
BI							
EE	0.618						
FC	0.69	0.735					
GPTU	0.812	0.52	0.502				
Hb	0.84	0.471	0.518	0.81			
HM	0.653	0.526	0.554	0.538	0.558		
PE	0.746	0.741	0.663	0.624	0.625	0.605	
SI	0.747	0.667	0.588	0.73	0.737	0.542	0.723

TABLE 4. Cross Loadings

	BI	EE	FC	GPTU	Hb	HM	PE	SI
BI1	**0.811**	0.391	0.462	0.601	0.698	0.478	0.512	0.606
BI2	**0.886**	0.48	0.484	0.588	0.639	0.554	0.617	0.533
BI3	**0.903**	0.541	0.557	0.665	0.64	0.503	0.592	0.575
BI4	**0.878**	0.474	0.504	0.629	0.582	0.508	0.569	0.542
EE1	0.312	**0.742**	0.419	0.174	0.217	0.248	0.403	0.313
EE2	0.378	**0.788**	0.43	0.348	0.284	0.397	0.506	0.409
EE3	0.474	**0.85**	0.545	0.372	0.346	0.411	0.527	0.46
EE4	0.513	**0.794**	0.473	0.479	0.404	0.377	0.543	0.57
FC1	0.478	0.437	**0.802**	0.272	0.297	0.333	0.442	0.301
FC2	0.433	0.468	**0.79**	0.35	0.371	0.293	0.373	0.366
FC3	0.462	0.455	**0.801**	0.276	0.279	0.345	0.438	0.355
FC4	0.424	0.483	**0.719**	0.38	0.376	0.474	0.45	0.465
GPTU1	0.705	0.403	0.382	**0.869**	0.721	0.435	0.527	0.549
GPTU2	0.535	0.317	0.26	**0.874**	0.636	0.352	0.388	0.476
GPTU3	0.599	0.453	0.42	**0.859**	0.604	0.437	0.471	0.567
Hb1	0.625	0.368	0.379	0.715	**0.901**	0.508	0.483	0.591
Hb2	0.654	0.359	0.391	0.668	**0.902**	0.43	0.475	0.536
Hb5	0.675	0.357	0.363	0.629	**0.853**	0.373	0.479	0.536
HM1	0.498	0.402	0.456	0.377	0.413	**0.898**	0.486	0.387
HM2	0.552	0.462	0.441	0.434	0.449	**0.936**	0.512	0.447
HM3	0.56	0.398	0.387	0.485	0.492	**0.914**	0.481	0.471
PE1	0.499	0.523	0.455	0.427	0.435	0.39	**0.817**	0.479
PE2	0.579	0.524	0.424	0.476	0.463	0.475	**0.858**	0.554
PE3	0.557	0.514	0.445	0.449	0.455	0.476	**0.857**	0.499
PE4	0.587	0.573	0.527	0.467	0.476	0.47	**0.85**	0.561
SI1	0.496	0.523	0.419	0.441	0.437	0.354	0.462	**0.825**
SI2	0.544	0.508	0.41	0.501	0.509	0.333	0.515	**0.829**
SI3	0.455	0.345	0.302	0.48	0.505	0.351	0.45	**0.78**
SI4	0.593	0.463	0.414	0.563	0.574	0.493	0.573	**0.815**

TABLE 5. Fornell – Larcker criteria

	BI	EE	FC	GPTU	H	HM	PE	SI
BI	0.871							
EE	0.543	0.795						
FC	0.577	0.593	0.779					
GPTU	0.714	0.454	0.412	0.867				
Hb	0.735	0.408	0.427	0.758	0.885			
HM	0.587	0.459	0.466	0.474	0.494	0.916		
PE	0.659	0.631	0.547	0.539	0.541	0.538	0.846	
SI	0.648	0.569	0.479	0.615	0.627	0.476	0.621	0.813

So, under the light of above calculation of Reliability measures of Cronbach Alpha, rho_A, Composite reliability and AVE and Validity measures of HTMT, Fornell- Larker criteria and cross loadings criteria proved the use of measurement model for further Hypotheses testing by using Structural model findings.

4.2. Structural Model

To assess the structural model of the current investigation, Standardized Root Mean Square Residual (SRMR) was calculated before determining the significance of the hypotheses under discussion (Habibi et al., 2023; Kono & Sato, 2022). The SRMR is a fit measurement used to prevent model misspecification in PLS-SEM (Magno et al., 2024). The ideal range for SRMR is 0.08 to 0.10, and the model can be educated to ensure a good fit. As shown in Table 6. The current model's SRMR value of 0.065 has proven to be a good fit. In addition, the squared Euclidean distance (d_ULS) and the geodesic distance (d_G) are provided, supporting the SRMR analysis (Table 6), in which both criteria have no specific measurement values; the values of d_ULS and d_G are 1.362 and 0.518, respectively (Hair et al., 2019).

The structural model used a boot-strapping procedure with 10,000 subsamples (Becker et al., 2022; Cheah et al., 2023). Path coefficient values (β, t, and p-values) were computed. Out of nine hypotheses, two are rejected (EE -> BI and FC -> GPTU), while the remaining seven are confirmed/supported.

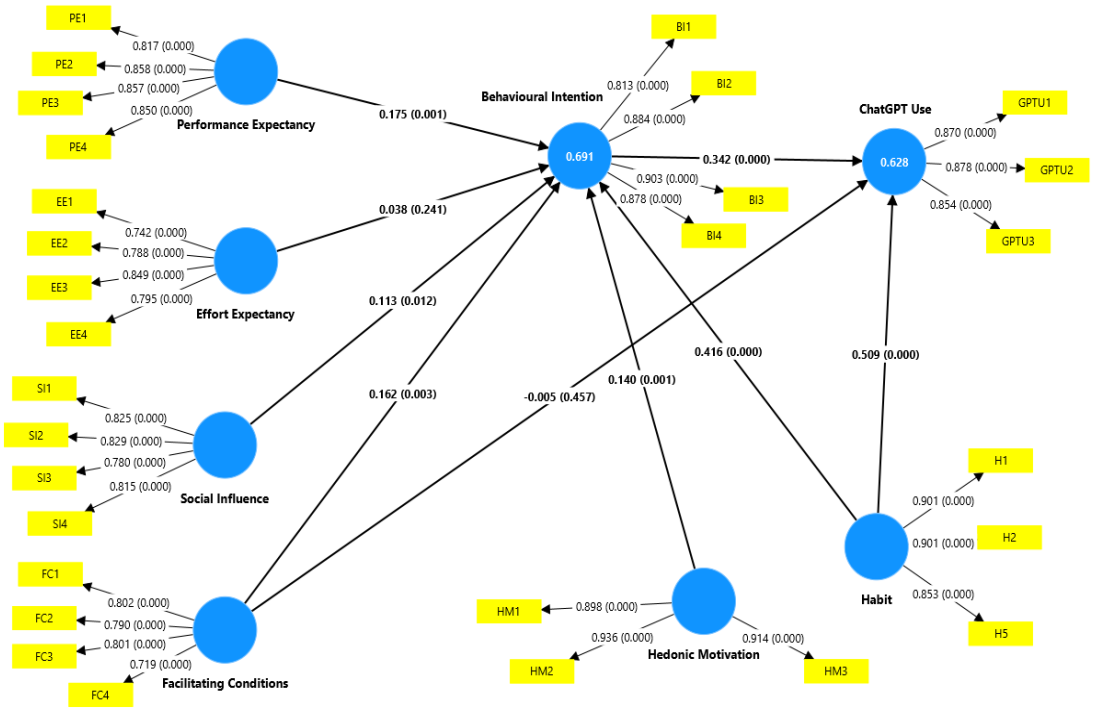
Four variables have a substantial impact on BI's use of ChatGPT in learning. In detail, FC is a significant predictor of BI followed by PE. The table 6 includes three values (β, t, and p), as well as the coefficient of determination (R2) and effect sizes (f2). R2 is the correlation of squares between dependent variables, which is used to assess how effectively an endogenous variable predicts external variables (Habibi et al.,2023). Figure 2 shows that all of the components (PE, EE, SI, HM, FC, and Hb) account for 69% of the variance within BI (R2 =0.691). Further variables (Hb, FC, and BI) explain 62.8% of the variation in the use of ChatGPT in learning among HEI students in Punjab, India.

R2 values are characterized as moderate (Cheah et al., 2023; Habibi et al., 2023; Hair et al., 2019; Kono & Sato, 2023). Along with this, effect sizes (f2) are calculated. Table 6 displays all f2 values for each relation in the model. The highest effect size emerges between H and GPTU (f2 =0.318), followed by the impact size between Hb and BI 0.3, while the smallest is seen in the association between EE and BI (f2 =0.002), followed by the effect between FC and BI (0.048). There is no effect size between FC and GPTU (f2 = 0.000).

TABLE 6. Structural model of factors affecting Indian HEI students' ChatGPT acceptance and use

Hypotheses	Relationship	Beta value	mean (M)	SD	T statistics	P values	Supported?	f2	Model fit
H1	PE -> BI	0.175	0.176	0.054	3.261	0.001	Supported	0.044	SRMR (0.065),
H2	EE -> BI	0.038	0.046	0.053	0.717	0.241	Not Supported	0.002	
H3	SI -> BI	0.113	0.115	0.05	2.247	0.012	Supported	0.018	
H4	HM -> BI	0.140	0.139	0.047	3.023	0.001	Supported	0.04	d_ULS(1.845),
H5	FC -> GPTU	-0.005	-0.004	0.042	0.043	0.457	Not Supported	0	
H6	FC -> BI	0.162	0.155	0.059	2.733	0.003	Supported	0.048	d_G(0.719)
H7	Hb -> BI	0.416	0.413	0.041	10.033	0	Supported	0.3	
H8	Hb -> GPTU	0.509	0.510	0.051	9.982	0	Supported	0.318	
H9	BI -> GPTU	0.342	0.341	0.056	6.166	0	Supported	0.118	

FIGURE 2. Final model investigating Indian HEI students' ChatGPT acceptance & use



(Performance Expectancy **[PE]**, Effort Expectancy **[EE]**, Social Influence **[SI]**, Facilitating Conditions **[FC]**, Hedonic Motivation **[HM]**, Habit **[Hb]**, Behavioral Intention **[BI]**, ChatGPT Use **[GPTU]**)

4.3. IPMA

The importance-performance map analysis (IPMA) expands on the findings of PLS-SEM by taking into consideration the performance of each component. As a result, conclusions may be reached on two dimensions (importance and performance), which is very useful for prioritizing managerial initiatives (Hair et al., 2021). Table 7 presents the needed results for the current investigation. The results demonstrate that H has a high impact on both BI (0.416) and GPTU (0.507). Following that, PE on BI had a beta value of 0.176. Higher values, ranging from 1 to 100, indicate higher performance. Table 7 reveals that EE has the highest performance (72.487), while GPTU has the lowest at 49.449.

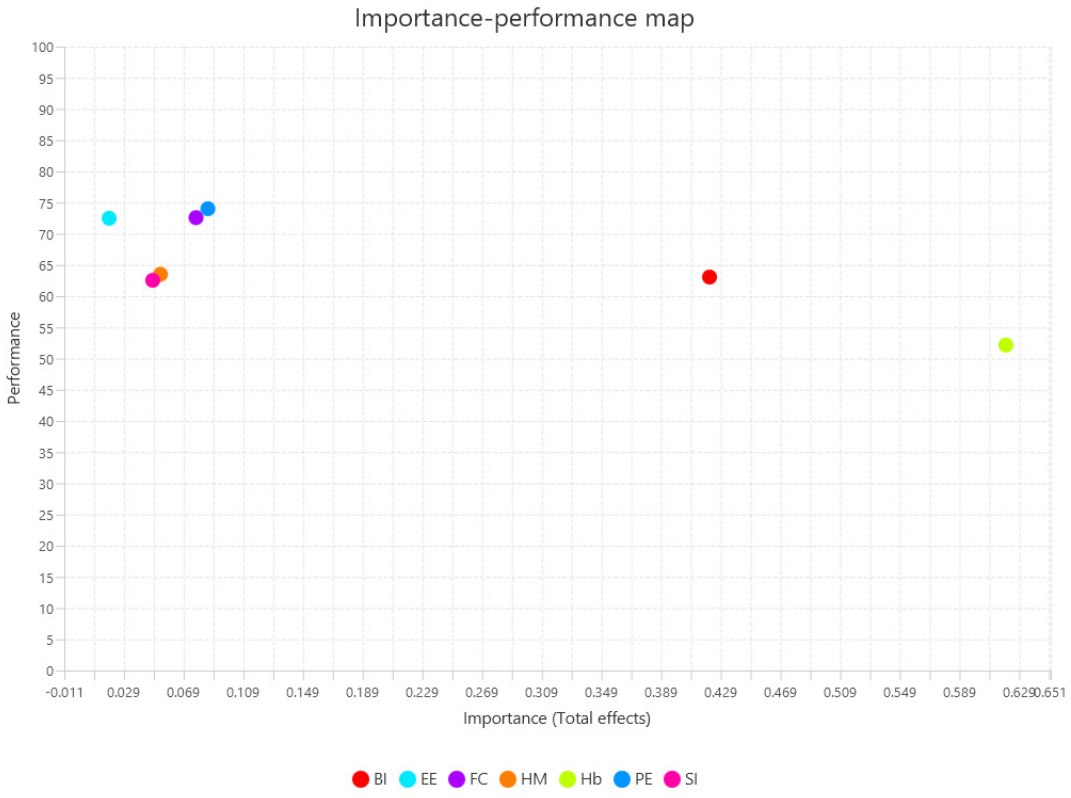
TABLE 7. IPMA results: Factors affecting ChatGPT BI and use

Performance		Importance	
		BI	GPTU
BI	63.066		0.422
EE	72.487	0.046	
FC	72.592	0.191	0.078
GPTU	49.449		
Hb	52.169	0.323	0.621
HM	63.513	0.128	
PE	74.018	0.203	
SI	62.538	0.115	

In Figure 3, IPMA (Importance-performance map), the x-axis reflects how important each factor is for the overall model. The Y-axis reflects the performance, showing how well each factor is rated by users in the context of ChatGPT usage among university students. From the map it can be deduced that PE (Performance Expectancy) and EE (Effort expectancy) were rated high which suggested that University students

find ChatGPT useful and easy to use. BI (Behaviour Intention) was crucial but did not perform as strongly as it could have, so it indicated that University curriculum can be planned so that it can improve and promote use of such tools in their learnings. In terms of SI (Social Influence) and Hb (Habit), these two constructs played a weaker role in affecting ChatGPT use among university students.

FIGURE 3. IPMA (Importance- Performance Map)



Aside from the responses to the questionnaire adopted from Habibi et al. (2023). Researchers also asked the students, "According to you, in which class/grade do students need training/guidance related to ChatGPT & Artificial Intelligence?" with four options: "6-8th Grade," "9-10 Garde," "11-12 Grade," and "Graduation." The percentage of the given options are as follows:

TABLE 8. Responses of HEIs students on "According to you, in which class/grade, students need the training/guidance related to ChatGPT & Artificial Intelligence"

Options	Frequency	%age
6-8th Grade	51	16%
9-10 Grade	86	27%
11-12 Grade	74	24%
Graduation	102	33%
Total	**313**	

The table 8 demonstrates that HEI students recognized the necessity to learn AI tools in schools. 67% of respondents prioritized learning these techniques while in school. In the Indian setting, as indicated by NEP (2020) under point no. 4.23, 4.24 at page no. 15, the addition of coding and artificial intelligence-related subjects at all levels, is a positive step. NEP 2020 also highlighted the same concern under point no. 23 i.e. "Technology Use and Integration", to establish an autonomous body, the National Educational Technology Forum (NETF) for the free exchange of ideas on the use of technology to enhance learning at both school and higher education level.

5. DISCUSSION

The structural model and IPMA analysis revealed that PE, SI, Hb, FC, and HM had substantial relationships with BI; Hb had a significant relationship with GPTU, and BI had a significant link with GPTU as well. However, there is no relevant link between EE and BI, nor between FC and GPTU. The findings revealed that Hb is the most important variable influencing BI to utilize ChatGPT while learning, as perceived by Punjab (India) higher education students.

Previous research has revealed the significance of Hb toward BI and Hb towards actual usage of ChatGPT (Cabrera-Sánchez et al., 2021; Chatterjee & Bhattacharjee, 2020; Fadzil, 2018; Foroughi et al., 2023; Gansser & Reich, 2021; Venkatesh, 2021). So, both of the significant outcomes (Hb -> BI and Hb -> GPTU) indicate that Indian Higher Education Students' habit of using ChatGPT influences their intention and use of ChatGPT in learning.

For the significance of SI in BI, the current results suggest that Social Influence (SI) plays an essential role in predicting BI when using ChatGPT. Furthermore, the data show that important others among Indian Higher Education Students have an impact on the BI of utilizing ChatGPT. Many prior studies have found similar results, including Alhwaiti (2023), Fadzil (2018), Guggemos et al. (2021), Habibi et al. (2023), Lin et al. (2022), and Mohd Rahim et al. (2022).

The significance of PE toward BI in the Indian Higher Education Context could be attributed to ChatGPT's function in enhancing learning tasks as well as the views given by ChatGPT on learning themes. Previous investigations have supported this conclusion (Alhwaiti, 2023; Andrews et al., 2021; Chatterjee & Bhattacharjee, 2020; Fadzil, 2018; Guggemos et al., 2020; Habibi et al., 2023; Lin et al., 2022; Mohd Rahim et al., 2022; Raffaghelli et al., 2022).

Vankatesh (2021) emphasized the relevance of FC in relation to BI and the actual application of ChatGPT, citing resource availability and technological system support. In the current study, FC was found to be a significant predictor of BI but not of ChatGPT use. As a result, it raises the worry that, while FC is an essential predictor for Indian students' behavioural intentions while utilizing ChatGPT, they have expressed concern about a lack of resources and support systems in higher education. Mohd Rahim et al. (2022) found similar results supporting FC to BI in the application of AI chatbots in Higher Education. This finding is consistent with the findings of Cabrera-Sanchez et al. (2021), who found that FC has a considerable influence on the actual use of AI but not on BI. As a result, the current findings highlight the support provided by technological systems to Indian higher education students.

In terms of the importance of HM in BI, the current findings are consistent with many earlier studies, including Azizi et al. (2020) (in a medical context), Arain et al. (2019) (mobile learning in higher education), and Habibi et al. (2023) (ChatGPT use in higher education). As a result, the findings reveal a particular inclination/motivation among Higher Education Students to use ChatGPT in their learning. So, students in Indian higher education like using ChatGPT.

In the instance when the relationship between EE and BI is not significant in the current data, Habibi et al. (2023) found similar results with EE. Some prior investigations found comparable results in the UTAUT environment (Andrews et al., 2021; Shivdas et al., 2020; Mohd Rahim et al., 2022). According to Habibi et al. (2023), the reason for this outcome could be that ChatGPT is an extension of other services such as Google and other technologies, and students in Higher Education perceive this as a normal modification in the list of already used technologies. In addition, BI demonstrated to be a strong predictor of actual ChatGPT use among Higher Education students. It supports the UTAUT2 model's claim that the more people intend to use new technology, the more and better they perform. So, it applies to ChatGPT and its application in learning (Habibi et al., 2023).

Practical implications

In terms of practical implications of this study, the results provide significant insights for universities, educators and policy makers to design the current and upcoming policy level, university level and classroom level changes in curriculum overall by considering use of ChatGPt and AI tools in higher education. This study highlights the insignificant influence of Facilitation Conditions on actual use of ChatGPT. It opens up the discussion for HEIs in India on what type of preparation is taking place at the policy and institutional levels in terms of training courses or technology support for students in relation to AI technologies. This study intended to spark a discussion about the intention and use of ChatGPT among Indian higher education students, notably those in Punjab. As NEP 2020 (recent education policy) highlighted the use and importance of AI in the educational experiences of students by introducing coding at grade 6; this study also made similar inputs that most of the students wanted the use of ChatGPT and AI tools at the school level to explore these new knowledge tools. As suggested by NEP 2020, the National Educational Technological Forum (NETF) will be an autonomous body for the leadership of education institutions, State and Central governments, and other stakeholders to decide the induction, deployment, and use of technology, by providing, the latest knowledge and research as well as the opportunity to consult and share best practices.

So, this study highlighted the importance of incorporating courses or modules in the higher education curriculum. With this, university students can understand the role and importance of ChatGPT and AI tools in the context of learning. Moreover, the facilitating conditions need to be taken care of, if universities want to improve the adoption and use of AI in higher education. Along with this, ethical considerations and professionalism need to be practiced with AI use and adoption among university students. Also, the results can also add to a better understanding of the impact and utility of future AI systems by focusing on the relevance of user intentions in aiming to maximize the potential of such technologies for learning.

HEIs need to review their policies related to academic honesty and integrity in relation to ChatGPT and other AI tools. It is imperative that some mentoring and support is provided to help staff and students

enhance their research technological expertise. Also, it is important that one helps students use ChatGPT as an adjunct to spontaneous human creativity and critical thinking.

Theoretical Implications

In terms of theoretical implications, this study provided the validation of the UTAUT2 model among university students in the context of Punjab, India. With 62.8% variance explained, this study gave compelling evidence in favour of the UTAUT2 model in the context of ChatGPT adoption and use among university students. In the context of a developing country, India can tackle the emergence of AI; the study's findings can be helpful in the adoption of AI technology. The results can be generalized to other developing countries considering the low level of awareness about AI tools.

6. LIMITATIONS AND FUTURE WORK

The current study on the use of ChatGPT in higher education, specifically among students in Punjab, India, highlights important insights but also faces some limitations. The research focuses on a single region-Punjab, limiting the generalizability of the findings across India's diverse socio-economic and educational contexts. Also, with a relatively small sample size and a focus only on students, the study does not account for teachers' perspectives, which are crucial in understanding the overall readiness for AI integration. Furthermore, concerns about academic integrity related to AI tools like ChatGPT were acknowledged but not explored in depth. The present study employed only quantitative methods for data analysis. For a more holistic perspective, research in the future could also use a combination of qualitative and quantitative methods; for eg., interviews of students could also be conducted to get a deeper understanding of their viewpoints regarding ChatGPT use.

Future research may focus on including other regions in the country, offering comparative studies to identify both common and unique factors affecting AI adoption. The inclusion of teachers' perspectives is essential, as they play a major role in implementing AI tools in the classroom. Additionally, multigroup analyses considering factors like age, gender, and cultural diversity could help provide more comprehensive insights. Researchers should also examine institutional readiness, policies, and support systems that facilitate AI integration, along with more focus on academic integrity and ethical use of AI. Expanding the scope of the UTAUT2 model and exploring technological infrastructure and readiness will further contribute to a more holistic understanding of AI's role in higher education.

7. CONCLUSION

The Horizon report (2023) in its Teaching and Learning Edition expressed concern for Higher Education Institutions about the need to plan for how to harness AI and its impending tools and services, such as ChatGPT, to improve efficiency and learning. However, this survey raises several issues about academic integrity, correctness, and cheating among higher education students. As a result, it encourages higher education

educators to be mindful of the introduction of AI tools such as ChatGPT into the lives of their students. So that curriculum and assessments may be created accordingly to incorporate AI tools such as ChatGPT into the learning process. Şimşek and Ateş (2022) concluded in their study that it is important to understand teachers' intentions to use Web 2.0 technologies in their respective courses which ultimately empowers students' efforts to use such tools in their own learning.

Moreover, present study results also highlight the insignificant influence of Facilitation Conditions on actual use of ChatGPT. It opens up the discussion for HEIs in India on what type of preparation is taking place at the policy and institutional levels in terms of training courses or technology support for students in relation to AI technologies. This study intended to spark a discussion about the intention and use of ChatGPT among Indian higher education students, notably those in Punjab.

8. REFERENCES

Abbas, J. (2018). Exploring the relationships among system quality, knowledge quality, and user satisfaction in the success of knowledge management systems: An empirical study. *Journal of Information & Knowledge Management, 17*(1), 1–21. https://doi.org/10.1142/S0219649218500079

Abu-Taieh, E. M., AlHadid, I., Abu-Tayeh, S., Masa'deh, R., Alkhawaldeh, R. S., Khwaldeh, S., & Alrowwad, A. (2022). Continued intention to use of M-banking in Jordan by integrating UTAUT, TPB, TAM, and service quality with ML. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(3), Article 120. https://doi.org/10.3390/joitmc8030120

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179-211. https://doi.org/10.1016/0749-5978(91)90020-t

Aldosari, S. A. (2020). The future of higher education in light of artificial intelligence transformations. *International Journal of Higher Education*, *9*(3), 145. https://doi.org/10.5430/ijhe.v9n3p145

Al-Fraihat, D., Joy, M., Masa'deh, R., & Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in Human Behavior*, *102*, 67-86. https://doi.org/10.1016/j.chb.2019.08.004

Alhwaiti, M. (2023). Acceptance of artificial intelligence application in the post-covid era and its impact on faculty members' occupational well-being and teaching self efficacy: A path analysis using the UTAUT 2 model. *Applied Artificial Intelligence*, *37*(1), Article 2175110 https://doi.org/10.1080/08839514.2023.2175110

Ali, M., Raza, S. A., Qazi, W., & Puah, C. (2018). Assessing E-learning system in higher education institutes. *Interactive Technology and Smart Education*, *15*(1), 59-78. https://doi.org/10.1108/itse-02-2017-0012

AlKursheh, T. (2024). Higher tertiary education perspectives: Evaluating the electronic assessment techniques of the blackboard platform for fairness and reliability. *Innoeduca. International Journal of Technology and Educational Innovation*, *10*(1), 144-165. https://doi.org/10.24310/ijtei.101.2024.17813

Al-Saedi, K., Al-Emran, M., Ramayah, T., & Abusham, E. (2020). Developing a general extended UTAUT model for M-payment adoption. *Technology in Society*, *62*, Article 101293. https://doi.org/10.1016/j.techsoc.2020.101293

Althunibat, A. (2015). Determining the factors influencing students' intention to use M-learning in Jordan higher education. *Computers in Human Behavior*, *52*, 65-71. https://doi.org/10.1016/j.chb.2015.05.046

Andrews, J. E., Ward, H., & Yoon, J. (2021). UTAUT as a model for understanding intention to adopt AI and related technologies among librarians. *The Journal of Academic Librarianship*, *47*(6), 102437. https://doi.org/10.1016/j.acalib.2021.102437

Arain, A. A., Hussain, Z., Rizvi, W. H., & Vighio, M. S. (2019). Extending UTAUT2 toward acceptance of mobile learning in the

context of higher education. *Universal Access in the Information Society*, *18*(3), 659-673. https://doi.org/10.1007/s10209-019-00685-8

Azizi, S. M., Roozbahani, N., & Khatony, A. (2020). Factors affecting the acceptance of blended learning in medical education: Application of UTAUT2 model. *BMC Medical Education*, *20*(1). https://doi.org/10.1186/s12909-020-02302-2

Becker, J., Cheah, J., Gholamzade, R., Ringle, C. M., & Sarstedt, M. (2022). PLS-SEM's most wanted guidance. *International Journal of Contemporary Hospitality Management*, *35*(1), 321-346. https://doi.org/10.1108/ijchm-04-2022-0474

Binyamin, S. S., & Zafar, B. A. (2021). Proposing a mobile apps acceptance model for users in the health area: A systematic literature review and meta-analysis. *Health Informatics Journal, 27*(1), Article 1460458220976737. https://doi.org/10.1177/1460458220976737

Bozorgkhou, N. (2015). An internet shopping user adoption model using an integrated TTF and UTAUT: Evidence from Iranian consumers. *Management Science Letters*, *5*(2), 199-204. https://doi.org/10.5267/j.msl.2014.12.017

Cabrera-Sánchez, J., Villarejo-Ramos, Á. F., Liébana-Cabanillas, F., & Shaikh, A. A. (2021). Identifying relevant segments of AI applications adopters – Expanding the UTAUT2's variables. *Telematics and Informatics*, *58*, Article 101529. https://doi.org/10.1016/j.tele.2020.101529

Camilleri, M. A. (2024). Evaluating service quality and performance of higher education institutions: A systematic review and a post-COVID-19 outlook. *Journal of Marketing for Higher Education, 34*(1), 1–24. https://doi.org/10.1080/08841241.2023.2165965

Chang, L., Wang, Y., Liu, J., Feng, Y., & Zhang, X. (2023). Study on factors influencing college students' digital academic reading behavior. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.1007247

Chatterjee, S., & Bhattacharjee, K. K. (2020). Adoption of artificial intelligence in higher education: A quantitative analysis using structural equation modelling. *Education and Information Technologies*, *25*(5), 3443-3463. https://doi.org/10.1007/s10639-020-10159-7

Cheah, J., Amaro, S., & Roldán, J. L. (2023). Multigroup analysis of more than two groups in PLS-SEM: A review, illustration, and recommendations. *Journal of Business Research*, *156*, Article 113539. https://doi.org/10.1016/j.jbusres.2022.113539

Choi, J. H., Hickman, K. E., Monahan, A., & Schwarcz, D. (2023). ChatGPT goes to law school. *71 Journal of Legal Education, 387*. https://doi.org/10.2139/ssrn.4335905

Dečman, M. (2015). Modeling the acceptance of E-learning in mandatory environments of higher education: The influence of previous education and gender. *Computers in Human Behavior*, *49*, 272-281. https://doi.org/10.1016/j.chb.2015.03.022

Dowling, M., & Lucey, B. (2023). ChatGPT for (finance) research: The banality of evil or the best of both worlds? *Finance Research Letters, 51*, 103362. https://doi.org/10.1016/j.frl.2023.103362

Duong, D., Nguyen, T., & Nguyen, A. (2023). The impact of ChatGPT on education: A study on the use of AI tools in academic settings. *Journal of Educational Technology Development and Exchange, 16*(2), 1-15. https://doi.org/10.18785/jetde.1602.01

Eke, D. O. (2023). ChatGPT and the rise of AI in creative writing: Implications for authors and publishers. *Publishing Research Quarterly, 39*(2), 203-217. https://doi.org/10.1007/s12109-023-09912-3

Fadzil, F. (2018). A study on factors affecting the behavioral intention to use mobile apps in Malaysia. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3090753

Farooq, M. S., Salam, M., Jaafar, N., Fayolle, A., Ayupp, K., Radovic-Markovic, M., & Sajid, A. (2017). Acceptance and use of lecture capture system (LCS) in executive business studies. *Interactive Technology and Smart Education*, *14*(4), 329-348. https://doi.org/10.1108/itse-06-2016-0015

Fornell, C., & Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research, 18*(1), 39-50.

Foroughi, B., Senali, M. G., Iranmanesh, M., Khanfar, A., Ghobakhloo, M., Annamalai, N., & Naghmeh-Abbaspour, B. (2023). Determinants of intention to use ChatGPT for educational purposes: Findings from PLS-SEM and fsQCA. *International Journal of Human–Computer Interaction*, *40*(17), 4501–4520. https://doi.org/10.1080/10447318.2023.2226495

Gansser, O. A., & Reich, C. S. (2021). A new acceptance model for artificial intelligence with extensions to UTAUT2: An empirical study in three segments of application. *Technology in Society*, *65*, Article 101535. https://doi.org/10.1016/j.techsoc.2021.101535

Guggemos, J., & Seufert, S. (2021). Teaching with and teaching about technology- Evidence for professional development of in-service teachers. *Computers in Human Behavior*, *115*, Article 106613. https://doi.org/10.1016/j.chb.2020.106613

Guggemos, J., Seufert, S., & Sonderegger, S. (2020). Humanoid robots in higher education: Evaluating the acceptance of Pepper in the context of an academic writing course using the UTAUT. *British Journal of Educational Technology*, *51*(5), 1864-1883. https://doi.org/10.1111/bjet.13006

Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence*, *5*, Article 100190. https://doi.org/10.1016/j.caeai.2023.100190

Habibi, A., Yaakob, M. F., & Sofwan, M. (2022). Student use of digital libraries during COVID-19: Structural equation modelling in Indonesian and Malaysian contexts. *The Electronic Library*, *40*(4), 472-485. https://doi.org/10.1108/el-12-2021-0212

Hair, J. F., Hult, G. T., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). An introduction to structural equation modeling. In *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R. Classroom Companion: Business* (pp. 1-29). Springer. https://doi.org/10.1007/978-3-030-80519-7_1

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2-24. https://doi.org/10.1108/ebr-11-2018-0203

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115-135. https://doi.org/10.1007/s11747-014-0403-8

Hoi, V. N. (2020). Understanding higher education learners' acceptance and use of mobile devices for language learning: A Rasch-based path modeling approach. *Computers & Education*, *146*, 103761. https://doi.org/10.1016/j.compedu.2019.103761

Huang, K. (2023). Artificial intelligence in education: A review of current developments and future prospects. *Educational Technology & Society, 26*(1), 105-117. https://doi.org/10.1007/s11423-023-10012-8

Im, I., Hong, S., & Kang, M. S. (2011). An international comparison of technology adoption: Testing the UTAUT model. *Information & Management, 48*(1), 1–8. https://doi.org/10.1016/j.im.2010.09.001

Kono, S., & Sato, M. (2022). The potentials of partial least squares structural equation modeling (PLS-SEM) in leisure research. *Journal of Leisure Research*, *54*(3), 309-329. https://doi.org/10.1080/00222216.2022.2066492

Kumar, J. A., & Bervell, B. (2019). Google Classroom for mobile learning in higher education: Modelling the initial perceptions of students. *Education and Information Technologies*, *24*(2), 1793-1817. https://doi.org/10.1007/s10639-018-09858-z

Kwateng, K. O., Atiemo, K. A., & Appiah, C. (2018). Acceptance and use of mobile banking: An application of UTAUT2. *Journal of Enterprise Information Management, 32*(1), 118–151. https://doi.org/10.1108/JEIM-03-2018-005

Lim, W. M., Kumar, S., & Ali, F. (2022). Understanding Generation Z: A review and future research agenda. *Journal of Business Research, 142*, 1-16. https://doi.org/10.1016/j.jbusres.2021.12.038

Lim, W. M., Kumar, S., & Ali, F. (2023). Exploring the potential of ChatGPT in higher education: A preliminary study. *Education and Information Technologies, 28*(3), 2501-2520. https://doi.org/10.1007/s10639-023-11021-2

Lin, H. C., Ho, C. F., & Yang, H. (2022). Understanding adoption of artificial intelligence-enabled language e-learning system: An empirical study of UTAUT model. *International Journal of Mobile Learning and Organisation*, *16*(1), 74. https://doi.org/10.1504/ijmlo.2022.119966

Lund, B. D., & Wang, T. (2023). ChatGPT in academia: Opportunities and challenges. *Journal of Academic Librarianship, 49*(1), Article 102123. https://doi.org/10.1016/j.acalib.2022.102123

Magno, F., Cassia, F., & Ringle, C. M. (2024). A brief review of partial least squares structural equation modeling (PLS-SEM) use in quality management studies. *The TQM Journal*, *36*(5), 1242-1251. https://doi.org/10.1108/tqm-06-2022-0197

Marto, A., Gonçalves, A., Martins, J., & Bessa, M. (2019). Applying UTAUT model for an acceptance study alluding the use of augmented reality in archaeological sites. In A. Tremeau, G. M. Farinella, & J. Braz (Eds.). *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (pp. 111-120). https://doi.org/10.5220/0007364101110120

Mogavi, O., Smith, J., & Lee, A. (2024). Balancing AI assistance and cognitive development in education. *Computers & Education, 182*, Article 104484. https://doi.org/10.1016/j.compedu.2022.104484

Mohd Rahim, N. I., A. Iahad, N., Yusof, A. F., & A. Al-Sharafi, M. (2022). AI-based chatbots adoption model for higher-education institutions: A hybrid PLS-SEM-Neural network modelling approach. *Sustainability*, *14*(19), Article 12726. https://doi.org/10.3390/su141912726

Moorthy, K., Chun T'ing, L., Ming, K. S., Ping, C. C., Ping, L. Y., Joe, L. Q., & Jie, W. Y. (2018). Behavioral intention to adopt digital library by the undergraduates. *International Information & Library Review*, *51*(2), 128-144. https://doi.org/10.1080/10572317.2018.1463049

Motaghian, H., Hassanzadeh, A., & Moghadam, D. K. (2013). Factors affecting university instructors' adoption of web-based learning systems: Case study of Iran. *Computers & Education*, *61*, 158-167. https://doi.org/10.1016/j.compedu.2012.09.016

Mustafa, S., Zhang, W., Shehzad, M. U., Anwar, A., & Rubakula, G. (2022). Does Health Consciousness Matter to Adopt New Technology? An integrated model of UTAUT2 with SEM-fsQCA approach. *Frontiers in Psychology*, *13*, Article 836194. https://doi.org/10.3389/fpsyg.2022.836194

Narayan, J., & Naidu, S. (2024). A new contextual and comprehensive application of the UTAUT2 model post-COVID-19 pandemic in higher education. *Higher Education Quarterly, 78*(1), 47–77. https://doi.org/10.1111/hequ.12441

Niloy, M. I., Rahman, M. M., & Islam, M. S. (2024). Factors influencing students' adoption of AI-based educational tools: A structural equation modeling approach. *Computers & Education, 182*, Article 104484. https://doi.org/10.1016/j.compedu.2022.104484

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator, 78*(1), 84-93. https://doi.org/10.1177/10776958221149577

Raffaghelli, J. E., Rodríguez, M. E., Guerrero-Roldán, A., & Bañeres, D. (2022). Applying the UTAUT model to explain the students' acceptance of an early warning system in higher education. *Computers & Education*, *182*, Article 104468. https://doi.org/10.1016/j.compedu.2022.104468

Ragheb, M. A., Tantawi, P., Farouk, N., & Hatata, A. (2022). Investigating the acceptance of applying chat-bot (Artificial intelligence) technology among higher education students in Egypt. *International Journal of Higher Education Management*, *8*(2), 1-13. https://doi.org/10.24052/ijhem/v08n02/art-1

Ramayah, T. J. F. H., Cheah, J., Chuah, F., Ting, H., & Memon, M. A. (2018). *Partial least squares structural equation modeling (PLS-SEM) using smart PLS 3.0: An updated guide and practical guide to statistical analysis* (2nd edition). Pearson

Ramu, P., Osman, M., Abdul Mutalib, N. A., Aljaberi, M. A., Lee, K., Lin, C., & Hamat, R. A. (2023). Validity and reliability of a questionnaire on the knowledge, attitudes, perceptions, and practices toward food poisoning among Malaysian secondary school students: A pilot study. *Healthcare*, *11*(6), Article 853. https://doi.org/10.3390/healthcare11060853

Raza, S. A., Qazi, W., Khan, K. A., & Salam, J. (2020a). Social isolation and acceptance of the learning management system (LMS) in the time of COVID-19 pandemic: An expansion of the UTAUT model. *Journal of Educational Computing Research*, *59*(2), 183-208. https://doi.org/10.1177/0735633120960421

Raza, S. A., Qazi, W., Umer, B., & Khan, K. A. (2020b). Influence of social networking sites on life satisfaction among university students: A mediating role of social benefit and social overload. *Health Education*, *120*(2), 141-164. https://doi.org/10.1108/he-07-2019-0034

Raza, S. A., Qazi, Z., Qazi, W., & Ahmed, M. (2021). E-learning in higher education during COVID-19: Evidence from blackboard

learning system. *Journal of Applied Research in Higher Education*, *14*(4), 1603-1622. https://doi.org/10.1108/jarhe-02-2021-0054

Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2018). Partial least squares structural equation modeling in HRM research. *The International Journal of Human Resource Management*, *31*(12), 1617-1643. https://doi.org/10.1080/09585192.2017.1416655

Salifu, I., Arthur, F., Arkorful, V., Abam Nortey, S., & Solomon Osei-Yaw, R. (2024). Economics students' behavioural intention and usage of ChatGPT in higher education: A hybrid structural equation modelling-artificial neural network approach. *Cogent Social Sciences*, *10*(1), Article 2300177. https://doi.org/10.1080/23311886.2023.2300177

Shivdas, A., Menon, D. G., & Nair, C. S. (2020). Antecedents of acceptance and use of a digital library system. *The Electronic Library*, *38*(1), 170-185. https://doi.org/10.1108/el-03-2019-0074

Şimşek, A. S., & Ateş, H. (2022). The extended technology acceptance model for Web 2.0 technologies in teaching. *Innoeduca. International Journal of Technology and Educational Innovation*, *8*(2), 165-183. https://doi.org/10.24310/innoeduca.2022.v8i2.15413

Soni, N., Kumar, S., & Singh, P. (2022). Generative AI: Transforming industries and society. *AI & Society, 37*(4), 1231-1245. https://doi.org/10.1007/s00146-021-01234-5

Strzelecki, A., & ElArabawy, S. (2024). Investigation of the moderation effect of gender and study level on the acceptance and use of generative AI by higher education students: Comparative evidence from Poland and Egypt. *British Journal of Educational Technology, 55*(3), 1209-1230. https://doi.org/10.1111/bjet.13425

Talan, T., Doğan, Y., & Kalinkara, Y. (2024). Digital natives' mobile learning adoption in terms of UTAUT-2 model: A structural equation model. *Innoeduca. International Journal of Technology and Educational Innovation*, *10*(1), 100-123. https://doi.org/10.24310/ijtei.101.2024.17440

Tamilmani, K., Rana, N. P., Prakasam, N., & Dwivedi, Y. K. (2019, June 1). The battle of Brain vs. Heart: A literature review and meta-analysis of "hedonic motivation" use in UTAUT2. *International Journal of Information Management, 46*, 222-235. https://doi.org/10.1016/j.ijinfomgt.2019.01.008

Thongsri, N., Shen, L., Bao, Y., & Alharbi, I. M. (2018). Integrating UTAUT and UGT to explain behavioural intention to use M-learning. *Journal of Systems and Information Technology, 20*(3), 278-297. https://doi.org/10.1108/jsit-11-2017-0107

Tseng, T. H., Lin, S., Wang, Y., & Liu, H. (2019). Investigating teachers' adoption of MOOCs: The perspective of UTAUT2. *Interactive Learning Environments*, *30*(4), 635-650. https://doi.org/10.1080/10494820.2019.1674888

Vaishya, R., Javaid, M., Khan, I. H., & Haleem, A. (2023). Artificial intelligence (AI) applications for COVID-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14*(4), 337-339. https://doi.org/10.1016/j.dsx.2020.04.012

Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, *27*(3), 425. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178. https://doi.org/10.2307/41410412

Venkatesh, V. (2021). Adoption and use of AI tools: A research agenda grounded in UTAUT. *Annals of Operations Research*, *308*(1-2), 641-652. https://doi.org/10.1007/s10479-020-03918-9

Vishwanath, A., & Goldhaber, G. M. (2003). An examination of the factors contributing to adoption decisions among late-diffused technology products. *New Media & Society*, *5*(4), 547-572. https://doi.org/10.1177/1461444480354005

Wang, W., & Wang, C. (2009). An empirical study of instructor adoption of web-based learning systems. *Computers & Education*, *53*(3), 761-774. https://doi.org/10.1016/j.compedu.2009.02.021

Wijaya, T., Cao, Y., Weinhandl, R., Yusron, E., & Lavicza, Z. (2022). Applying the UTAUT model to understand factors affecting micro-lecture usage by mathematics teachers in China. *Mathematics*, *10*(7), Article 1008. https://doi.org/10.3390/math10071008

Yeop, M. A., Wong, K. T., & Noh, N. M. (2016). Blended learning in selected journals: A content analysis using the complex adaptive blended learning systems. *International Journal of Instructional Technology and Distance Learning, 13*(10), 47–58.

Zacharis, G., & Nikolopoulou, K. (2022). Factors predicting University students' behavioral intention to use eLearning platforms in the post-pandemic normal: An UTAUT2 approach with 'Learning value'. *Education and Information Technologies*, *27*(9), 12065-12082. https://doi.org/10.1007/s10639-022-11116-2

Zwain, A. A. (2019). Technological innovativeness and information quality as neoteric predictors of users' acceptance of learning management system. *Interactive Technology and Smart Education*, *16*(3), 239-254. https://doi.org/10.1108/itse-09-2018-006

9. APPENDIX

Instrument (Habibi et al., 2023- English Version)

ENGLISH

This survey looks at your perspective on the use of ChatGPT in learning

- Gender [Male/female]
- Age
- Do you use ChatGPT

If yes, please continue ☐

UTAUT

Performance Expectancy [PE], PE1-PE4

1. ChatGPT is useful to carry out my tasks.
2. Using ChatGPT would increase the efficiency of my work.
3. Using ChatGPT would improve the quality of my tasks.
4. Using ChatGPT would allow me to have more convenient at work.

Effort Expectancy [EE], EE1-EE4

1. It is easy to enter in the ChatGPT page.
2. My interactions with my mobile phone and transaction terminals when using ChatGPT are clear and understandable.
3. I find it easy to use ChatGPT.
4. It is easy for me to become skilful at using ChatGPT.

Social Influence [SI], SI1-SI3

1. My friends think that I should use ChatGPT.
2. My family think that I should use ChatGPT.
3. People who influence my behaviour use ChatGPT.
4. The use of ChatGPT gives me professional status.

Facilitating Conditions [FC], FC1-FC4

1. I have the necessary resources (laptop, internet connection, mobile, desktop, etc.) to use ChatGPT.
2. I have the necessary knowledge to use ChatGPT.
3. The ChatGPT is compatible with the existing technology (like windows/mac for laptop, android/iOS for mobile, etc.) that I use.
4. I can get help from others when I have difficulty in using the ChatGPT.

Hedonic Motivation [HM], HM1-HM3

1. Using ChatGPT system is fun.
2. Using ChatGPT system is enjoyable.
3. Using ChatGPT system is very entertaining.

Habit [Hb], H1-H5

1. The use of ChatGPT has become a habit for me.
2. I don't even think twice before using the ChatGPT.
3. Using the ChatGPT has become natural to me.
4. Using the ChatGPT has become automatic for me.
5. When faced with research or assignments, using the ChatGPT is an obvious choice.

Behavioral Intention [BI], BI1-BI4

1. I intend to use the ChatGPT.
2. I like using the ChatGPT.
3. I plan to continue to use the ChatGPT.
4. I will recommend my friends to use the ChatGPT.

ChatGPT Use, GPTU1-GPTU3

1. I use frequently use ChatGPT.
2. I depend on ChatGPT for learning.
3. I use all functions in ChatGPT.

According to you, in which class/grade, students need the training/guidance related to ChatGPT & Artificial Intelligence

1. 6-8th Grade.
2. 9-10th Grade.
3. 11-12th Grade.
4. Graduation.

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 29-46 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20547

Digital Entrepreneurship Education: a challenge for competency-based training in universities

Educación en emprendimiento digital: un reto para la formación por competencias en las universidades

RECEIVED 24/09/2024 ACCEPTED 08/10/2024 PUBLISHED 01/06/2025

M. Paz Prendes-Espinosa
ORCID 0000-0001-8375-5983
Universidad de Murcia, España
pazprend@um.es

Víctor González-Calatayud
ORCID 0000-0002-6835-0264
Universidad de Murcia, España
victor.gonzalez@um.es

José Luis Serrano
ORCID 0000-0003-2359-959X
Universidad de Murcia, España
jl.serranosanchez@um.es

ABSTRACT

This article focuses on the analysis of digital entrepreneurship as a transversal competence for university students. We take as a starting point the EmDigital (Digital Entrepreneurship) model that includes four areas of competence. The aim of the study was to analyse the digital entrepreneurship competence of final-year university students. It is a mixed research that includes a survey of university students (representative cluster sampling) and interviews with successful digital entrepreneurs (non-probabilistic convenience sampling). Two ad hoc instruments have been used and previously validated. The sample of final year undergraduate students who participated in the survey was 1108 and the sample of successful entrepreneurs interviewed was 20. Our main results show that although students perceive themselves as having average competences for digital entrepreneurship, there are significant differences depending on the branch of knowledge of their university studies. Moreover, there are some differences in terms of gender. The interviewees agree on the importance of training university students for digital entrepreneurship skills and the need to consider this competence as basic skill in formal education.

KEYWORDS Digital entrepreneurship; competency; higher education; mixed research; field of knowledge; gender.

RESUMEN

Este artículo se centra en el análisis del emprendimiento digital como competencia transversal para estudiantes universitarios. Se toma como punto de partida el modelo EmDigital (Emprendimiento Digital) que se compone de cuatro áreas competenciales. El objetivo del estudio fue analizar la competencia de emprendimiento digital de alumnado universitario

de último curso. Se trata de una investigación mixta que incluye una encuesta a estudiantes universitarios (muestreo representativo por conglomerados) y entrevistas a emprendedores digitales de éxito (muestreo no probabilístico de conveniencia). Se han utilizado dos instrumentos ad hoc previamente validados. La muestra de estudiantes universitarios de último curso que participaron en la encuesta fue de 1108 y la muestra de emprendedores de éxito entrevistados fue de 20. Nuestros principales resultados muestran que, aunque los estudiantes se perciben a sí mismos con competencias medias para el emprendimiento digital, existen diferencias significativas en función de la rama de conocimiento de sus estudios universitarios. Además, existen algunas diferencias en función del género. Los entrevistados coinciden en la importancia de formar a los estudiantes universitarios en competencias de emprendimiento digital y en la necesidad de considerar esta competencia como básica en la educación formal.

PALABRAS CLAVE Emprendimiento digital; competencia; universidad; investigación mixta; área de conocimiento; género.

1. INTRODUCTION

Entrepreneurship in the 21st century cannot be understood without its association with digitalization, as they end up being two sides of the same coin, but not all authors agree on how to understand this concept of digital entrepreneurship. On the one hand, some authors associate digital entrepreneurship with the idea of profitability and business from a strictly business perspective, as is the case with the approach of Satalkina and Steiner (2020) or Bogdanowicz (2015).

However, from our perspective and according to other different approaches, the business focus is not so relevant, but the focus must be on competence and its indicators, which allows us to better address the education and training of a digital entrepreneur. This is the perspective of Nambisan (2017), who considers digital entrepreneurship as a concept that involves the intersection between entrepreneurship and digital technologies in today's digital society. In this line we also find Kollmann (2009), Lorenzo (2012) or Omar et al. (2019). All of them serve as references to understand digital entrepreneurship as the ability to develop innovative ideas that, using digital technologies, involve value creation in a sustainable context.

On the other hand, Giones and Brem (2017) differentiate between technological entrepreneurship (creation of innovative products in the field of technologies), digital technological entrepreneurship (creation of products specifically supported by digital technologies such as the internet of things or smart devices) and digital entrepreneurship, understood as innovation (new products or services) supported by the internet, such as new products in the cloud, cloud computing, data mining, etc.

In addition to these concepts, we can consider "digital academic entrepreneurship" (Secundo et al., 2020), which is defined as the use of digital technologies to promote entrepreneurship in an academic context, taking into account four different possibilities: digital technologies for entrepreneurship training, the maker movement, digital technologies as tools to discover entrepreneurship opportunities (linked to start-ups or patents and innovative ecosystems) and finally, the promotion of entrepreneurship skills in digital university spaces linked to innovative projects or courses aimed at this purpose.

1.1. Entrepreneurship education

It is understood, as a starting point, that the capacity for entrepreneurship is not innate, but is developed and learnt, so it is relevant to understand the need of a formal entrepreneurship education in universities. Although some studies recognize certain characteristics in the personality of entrepreneurs, these are dimensions that can be worked on in formal training processes Fossatti et al. (2023) consider entrepreneurship education in higher education institutions as one of the main factors to improve employability of students, moreover taking into account the need to align the competences provided by institutions with the skills demanded by the employers.

Gutiérrez-Mora et al. (2021) point out that the most necessary trait of the entrepreneur is self-confidence and, on a second level, other traits such as the ability to detect opportunities, risk tolerance, initiative, responsibility, authority, the need for achievement, personal values such as perseverance, motivation and, also very important, the ability to face up to problems.

Beyond trying to promote conditions (political, social or economic) for the design of entrepreneurial proposals and initiatives, we must think about the importance of training our students from the initial stages of the system to become entrepreneurs. Entrepreneurship education is a relatively new approach. Ruiz (2012) indicated that there was hardly any reference to this specific training in academic university education. However, the European Union has now included training in entrepreneurial skills as one of the fundamental elements to face the economic, social and cultural changes of the present and the future (European Commission, 2018) and universities must create the optimal conditions to promote this training (González-Calatayud et al., 2018; Mababu, 2017; Ratten & Usmanji, 2021). As Standing and Mattsson (2016) point out, this is not an easy goal, since there are certain gaps in knowledge about how digital entrepreneurs identify business opportunities and conceptualize them into a business model, an aspect that we will explore in more depth in our empirical research. It will be necessary to "find innovative teaching methods and develop critical thinking skills" (Ratten & Usmanji, 2021, p. 7).

Moreover, recent data show that the level of entrepreneurship in Spain, although stable, is far behind many other countries (GEM, 2020). In Spain, in general, initiatives on entrepreneurship training are very limited and of little depth compared to other countries in our European and international environment, where the relevance of such training is recognized. It is true that the current education law (BOE, 2020) points to the need to encourage and develop the entrepreneurial spirit. Similarly, in the Higher Education system, entrepreneurial training is also a pending issue. It is important to understand that training in entrepreneurial competences requires addressing all its dimensions and indicators, thus abandoning the business approach that has traditionally been associated with entrepreneurship and focusing on a more globalizing approach that allows us not only to analyse the competence, but also to design successful training proposals.

1.2. Digital entrepreneurship models

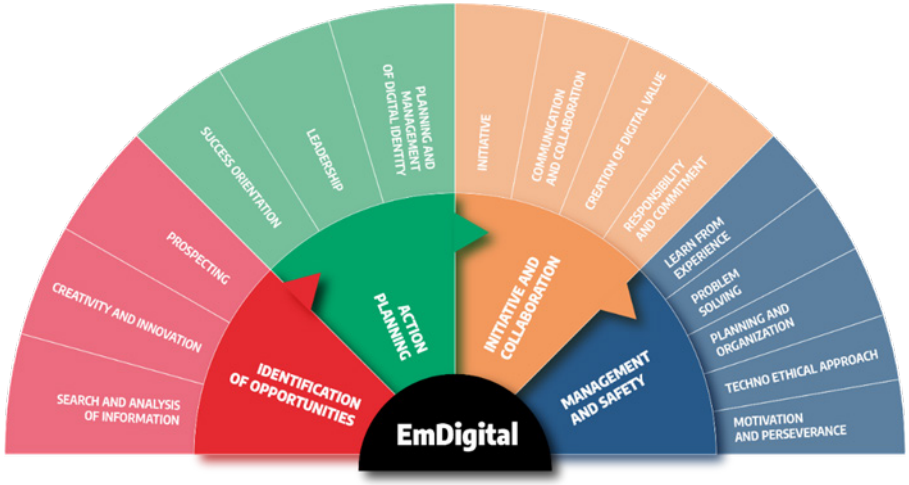
The research shows us several digital entrepreneurship models. Basly and Hammouda (2020) highlight that one of the main needs of digital entrepreneurship is to have a theoretical model that captures all the dimensions necessary for its analysis and for understanding the needs of digital entrepreneurs. Some models

analyse the competence of entrepreneurship (not digital entrepreneurship), although they propose among their dimensions some indicators that can be considered linked to digital competences. This is the case of the EntreComp model (Bacigalupo et al., 2016; Bacigalupo, 2022; McCallum et al. 2018), with three competence areas and 15 sub-competences. In the university context, the HEInnovate model, developed in 2012 between the OECD and the European Commission, is worth mentioning; it is a self-reflection tool for higher education institutions on their innovative potential. Other proposals include the model of Jardim (2021), who builds a reference framework that includes innovation, value creation and effective communication, based on the analysis of contexts and problems, as well as challenges and opportunities.

However, we find models that specifically break down the competence of digital entrepreneurship. Thus Carreón (2014) identifies five areas: basic digital knowledge and knowledge of the ICT market; digital business; access to finance and investment; digital skills and digital leadership; and finally, entrepreneurial culture. On his hand, Cruz García (2016) also proposes four areas of interest: digital agenda, teaching-learning, digital skills and employment. In a previous work, Cruz García (2015) conducts a documentary review and concludes that perceptions of compatibility, usefulness and ease of use are essential to explain the process of adoption and use of technologies for entrepreneurship. For their part, Pérez et al. (2016) design a model of digital entrepreneurship that does not rely on competences, but describes a sequence in which values, beliefs and perceptions are the basis on which people will develop their attitude and knowledge to tackle entrepreneurial projects.

And finally, we need to explain the EmDigital model (see Figure 1) on which our research is based and whose design has involved a complex process of qualitative research supported by documentary analysis techniques, content analysis, focus group and expert judgement through the Delphi method (García-Tudela et al., 2020; Prendes-Espinosa, 2022; Prendes-Espinosa & García-Tudela, 2020). The result of this complex process is the EmDigital model of digital entrepreneurship for university students, which includes 4 competence areas with 15 sub-competences and 45 indicators (García-Tudela et al., 2020; Prendes-Espinosa et al., 2021; Prendes-Espinosa & García-Tudela, 2020). The EmDigital areas are: identification of opportunities; action planning; implementation and collaboration; and finally, management and security.

FIGURE 1. EmDigital model about digital entrepreneurship (Prendes et al., 2021, p. 9)



2. RESEARCH PROBLEM AND OBJECTIVES

Based on the analysis of previous work on digital entrepreneurship, we have focused on the research problem about how the digital entrepreneurship competence is being worked on in Spanish universities. This problem has been specified in the following general and specific research objectives (Table 1).

TABLE 1. Research objectives

General Objectives	Specific objectives
To analyse the self-perception of digital entrepreneurship competence of Spanish university students	To describe the self-perception about digital entrepreneurship of Spanish university students in relation to every dimension of EmDigital model. Compare data on self-perception in each dimension of the EmDigital model according to gender and area of knowledge. To propose recommendations about training programs for university students to promote digital entrepreneurship.
To identify the strategies used by successful digital entrepreneurs.	To analyse the perception of successful digital entrepreneurs about the main factors to promote entrepreneurship initiatives. To identify the main strategies to promote successful practices in relation to the processes of opportunity identification, planning and implementation of the digital entrepreneurial action.

3. MATERIALS AND METHOD

3.1. Research design

Based on our model EmDigital (Figure 1), we have designed our research using a mixed method which combines quantitative and qualitative techniques. A quantitative research phase was carried out using the survey technique (with an ad hoc questionnaire validated with diverse techniques like a focus group, expert judgement, cognitive interviews and a pilot study) and a qualitative phase using interviews (validated with a focus group) with successful digital entrepreneurs who have been considered as examples of good practices in digital entrepreneurship.

3.2. Data collection tools

A questionnaire was designed and validated. Based on the theoretical model of digital entrepreneurship (EmDigital), a questionnaire was created that has gone through several validation phases: expert judgement (8 experts in the field of educational technology and entrepreneurship validated the instrument), cognitive interviews (19 interviews) and a pilot test. After all this process, the questionnaire is composed of 55 Likert-type items with 5 levels - 1 strongly disagree and 5 strongly agree - and a “no answer” option, which fall into 4 dimensions: Identification of opportunities (8 items), Action plan (11 items), Interactive and Collaboration (15 item), and Management and Safety (21 item). The CFA confirms a 4-dimensional structure that explains 47% of the total variance. The Bartlett’s test obtained a value of $\chi 2=25,265.83$, with 1,485

degrees of freedom and a p-value of $p<0.05$. The KMO coefficients were: 0.85, 0.86, 0.92, and 0.95. The overall model fit was excellent, with an RMSR value of 0.03. Cronbach's alpha of the overall questionnaire was .966. The reliability per dimension was as follows: 1. Identification of opportunities .83; 2. Action planning .84; 3. Implementation and collaboration .90; and 4. Management and safety .94. The questionnaire can be found at the following link in its Spanish version: https://acortar.link/8i9a4t

In the qualitative phase, a semi-structured interview was designed based on the analysis of good practices and previous work on successful digital entrepreneurship experiences. The interview was validated using the expert judgement method. Subsequently, the interviews were conducted by videoconference, as they were carried out during the Covid pandemic.

The main purpose of the interview was to determine the factors that influence the success of an entrepreneur in the digital world. To achieve this purpose and guide the content of the items, the following interview objectives were established:

- To know the keys to idea development and opportunity identification.
- To know the main aspects of entrepreneurial action planning and business design.
- To know the main aspects of the implementation of entrepreneurial action and the collaboration with other agents.
- To address global aspects of the keys to success.

Documentary review and collaborative procedure were the main methods followed to design the semi-structured interview. This initial phase was used to design the indicators of good practices in digital entrepreneurship. The interview script was validated using the expert judgment technique. All members of the EmDigital Project research team participated in a two-phase validation sequence. The final version of this instrument, after applying adjustments during the validation process, can be consulted in Prendes-Espinosa et al. (2021).

3.3. Context of the study and research participants

For the survey we have used the population of university students in the Region of Murcia (Spain). The questionnaire was applied to a sample representative of 1108 final-year university students selected by clusters. A sample of 20 digital entrepreneurs was chosen (non-probabilistic discretionary sampling) to be interviewed.

The sample selection for the questionnaire was carried out using a stratified random cluster sampling technique. Universities and branches of knowledge were used as the strata for the distribution. Subsequently, the degrees included in the branch and the final year under-graduate classrooms were used as the clusters. Firstly, the minimum representative sample was established with a confidence level of 95% and a sampling error of 0.03, the minimum sample being 1032.07 for a population of 31951 students.

TABLE 2. Distribution of participants by field of knowledge

Field of knowledge	Students	Required sample	Participants
Arts and Humanities	3317	107	157
Social and Legal Sciences	14520	469	503
Health Sciences	5957	192	152
Engineering and Architecture	1370	44	39
Sciences	2596	84	140
Engineering and Architecture	3778	122	117

The final sample of participants consisted of a total of 1108 final year undergraduate students from the public universities of the Region of Murcia. Specifically, 64.17% were women, with a mean age of 22.81 (SD= 3.28). The final distribution by branches of knowledge and University according to the total population is shown in Table 2.

For the selection of the participants to be interviewed, firstly it was necessary to determine the characteristics and requirements necessary to define the so-called good practices of digital entrepreneurship. The main sources used were the next: EmDigital model (see Figure 1) and the Strategic Tree (Ogel & Castillo, 2012), a visual management tool oriented towards strategic thinking for entrepreneurs.

A non-probabilistic convenience sampling was used to select the participants. In the first round, 45 cases were selected after applying some indicators of good practices in digital entrepreneurship, these were the next:

- A prize, award, or recognition of success.
- Person who has funding or is hosted in development and innovation centers.
- Clearly defines mission motivations (what for) and opportunities.
- It contemplates the customer in the development and distribution of the product.
- A clear and accessible digital corporate identity.
- Develops and uses digital spaces to foster collaboration and participation of different agents.
- Develops, uses and disseminates digital content.

However, it was not possible to interview all those selected entrepreneurs due to the complications that many of them experienced during the Covid-19 crisis. Others declined the possibility of being interviewed for personal reasons or time constraints. Finally, 20 people were interviewed, all of them meeting the criteria previously defined to be considered successful digital entrepreneurs.

3.4. Data collection process

To collect the quantitative data from the survey, the questionnaire was used in an online form, but the information was collected in person in the university classrooms. After the random sampling process, the researchers wrote via email to teachers involved to ask for permission. After having the teachers' permission, researchers went in person to the classrooms and its purpose was explained before applying the

questionnaire. A QR code was used to facilitate access to the online questionnaire, which included an informed consent check box at the beginning. Those who consented to answer the questionnaire used the link provided at the time, so the research complies with the ethical requirements.

To access successful entrepreneurs, after searching for those chosen for having received official and public awards or recognition, they were contacted by e-mail. An authorization (informed consent) was requested to proceed with the interviews and their recording. All interviews were conducted virtually, audio-recorded and conducted by a member of the research team. Data was collected in 2020/21 and a protocol was designed for the interviewers, all of whom were members of the project's research team.

The procedure and instruments were approved by the Ethic Committee of the University of Murcia taking into account the data protection of participants too.

3.5. Data analysis

The results of the questionnaire have been analyzed using SPSS (version 24). The mean and frequency were used for descriptive analysis of the data. Non-parametric techniques were used to analyze possible differences according to sex and branch of knowledge, as these are the most robust tests for ordinal variables. Thus, the Mann-Whitney U test was used for crosses with two-level factors, and the Kruskal-Wallys H test was used for crosses with factors of three or more levels. When significant differences are found, the effect size is also shown, using Cohen's d (0.2 small effect, 0.5 medium and 0.8 large) for two-group crosses and eta-squared for crosses with three or more levels (0.01 small effect, 0.06 medium and 0.14 large).

About the qualitative information, the categorization strategy used was the coding of text fragments and the use of dimensions of analysis (Maxwell & Miller, 2008). Four cycles were followed to establish the final codebook combining inductive and deductive methods. As a first step in the data analysis process, the researchers responsible for coordinating the interview design and validation task created an initial outline of initial categories and subcategories. Subsequently, the transcripts were actively read, noting tentative ideas about the functioning of the categories and subcategories already created and their relationships according to the data. After updating the initial list of categories and codes, two members of the research team analyzed 5 interviews based on this list. In addition, codes were created for each of the categories and subcategories and examples of the interviews that corresponded to each code were indicated. After sharing the analysis of the 5 interviews, the researchers designed a new codebook, which was then refined following an emergent analysis of the data. After completing the data analysis with the qualitative analysis tool Nvivo, a codebook with 5 major categories, 16 subcategories and 66 codes was obtained.

4. RESULTS

4.1. Quantitative method: questionnaire for university students

Following the EmDigital model, the results of the quantitative part are shown in relation to the four dimensions that constitute the EmDigital model explained above: identification of opportunities; action planning; initiative and collaboration; and finally, management and security.

4.1.1. First dimension: identification of opportunities

The dimension of identification of opportunity is composed of three competences: information search and analysis, creativity and innovation, and prospecting. The following Table 3 shows the descriptive values of these competences and of the dimension.

TABLE 3. Descriptive data for the dimension Identification of opportunities

	Minimum	Maximum	Mean	SD
Information search and analysis	3	20	13.65	3.01
Creativity and innovation	1	10	6.28	1.98
Prospecting	1	10	7.6	1.61
Identification of opportunities	8	40	27.44	5.59

If we look at the differences according to sex, we find significant differences, with higher scores for men (Md= 28, Range=32) than for women (Mdn=27, Range=32), U= 123686, p=.001, PSest= 0.21. With regard to the branch of knowledge in this dimension, we could see how the differences between the branches are pronounced, especially be-tween the Science branch (M= 24,9) with Engineering and Architecture (M= 28,63, p< .001) IC95% [2.05,5.42] as well as with Social and Legal Sciences (M= 28,06, p=.002) IC95% [1.77,4.56], with this difference being significant, H(4)= 49.98, p< .001.

4.1.2. Second dimension: action planning

This dimension is composed of the competences: success orientation, leadership and planning and management of digital identity. Table 4 below shows the descriptive values of the dimension.

TABLE 4. Descriptive data for the dimension Action planning

	Minimum	Maximum	Mean	SD
Success orientation	5	25	21.56	3.4
Leadership	3	15	11.3	2.27
Planning and management of digital identity	1	15	11.19	2.67
Action planning	5	55	44	6.59

4.1.3. Third dimension: initiative and collaboration

The third dimension is composed of four competences: initiative, communication and collaboration, creation of digital value and responsibility and commitment. The descriptive data for the dimension and these competences can be found in Table 5 below.

TABLE 5. Descriptive data for the dimension Initiative and collaboration

	Minimum	Maximum	Mean	SD
Initiative	1	10	6.09	1.97
Communication and collaboration	2	30	21	5.37
Creation of digital value	1	15	10.37	2.84
Responsibility and commitment	2	20	15.14	3.39
Initiative and collaboration	10	75	52.5	11

In this dimension, the Mann Whitney U-test determines the acceptance of the null hypothesis, so no significant differences were found between males and females. The Kruskal Wallis test indicates that the null hypothesis must be rejected and that there are differences between the different branches of knowledge, H(4)= 46.07, p< .001. Using the Games-Howell post hoc test, it can be verified how the branch of Sciences (M= 47.51) and Health Sciences (50.91) obtained a significantly lower score than Social and Legal Sciences (54.24, p<.01) IC95% [4.07,9.40] and (p<.01) IC95% [0.39,6.28].

4.1.4. Four dimension: management and safety

The last dimension is composed of up to 5 competences: learn from experience, problem solving, planning and organization, techno ethical approach and motivation and perseverance. As before, the following Table 6 shows the descriptive data.

TABLE 6. Descriptive data for the dimension Management and safety

	Minimum	Maximum	Mean	SD
Learn from experience	2	20	14.13	3.57
Problem solving	3	25	17.24	4.15
Planning and organization	4	20	13.76	3.3
Techno ethical approach	1	20	13.96	3.49
Motivation and perseverance	1	20	13.53	3.67
Management and safety	14	105	72.41	15.99

In this dimension, the Mann Whitney U-test determines the acceptance of the null hypothesis, so no significant differences were found between males and females. The Kruskal Wallis test indicates that the null hypothesis must be rejected and that there are differences between the different branches of knowledge, H(4)= 84.23, p< .001. Using the Games-Howell post hoc test, it was verified how the Science branch (M=61.57) obtained significantly lower scores than the rest of the branches: Arts and Humanities (M=70.09, p<.001) IC95% [3.06,13.99], Health Sciences (M=71.58, p<.001) IC95% [4.86,15.17], Social and Legal Sciences (M=75.26, p<.001) IC95% [9.62,17.77], and Engineering and Architecture (M=75.6, p<.001) IC95% [9.31,18.76].

4.2. Qualitative results: interviews to successful digital entrepreneurs

Knowing the keys to the development of ideas and identification was one of the objectives of the interview conducted with the successful digital entrepreneurs who participated in the research. The essential point analyzed was whether the entrepreneurs really make a prior reflection on their experiences, beliefs, values, and training available to them and if all of the above connects with a purpose. The results show that in only two cases did they carry out a prior reflection of this initial step. The rest of the participants commented on superficial reflections on this supposedly essential step.

Yes, but I started with the idea of the type of business I wanted to create based on the kind of lifestyle I wanted, not the idea of the product itself. I wanted to work online. I wanted to work remotely, to have the freedom to work here and there. So, I thought of creating a product that would allow me to do that. I didn't really care about the specific product idea.

It is also true that once they have launched their entrepreneurial project, it seems they have a clearer understanding of their purpose. In other words, it may not be clear at the beginning, but during the development process, they become more aware of the purpose the product or idea they are working on serves. At least that's what six cases indicated.

Well, my ideal mission... I think it has always been about reaching people and helping them. Besides my studies... my interest lies in personal growth, motivation, coaching. So, if I do something, I would like it to help others in some way... to serve a purpose beyond being just a consumer product or something that passes through your life without any significance.

Understanding the key aspects of entrepreneurial action planning is the second objective of the interview. More than half of the participants implement strategies to organize and differentiate between the urgent and the important. It is surprising that many entrepreneurs do not apply this productivity strategy widely used in the entrepreneurial world when it comes to organization. Nevertheless, they recognize the need for it.

I have the theory, but in daily practice, I'm kind of jumping from one thing to another. Just yesterday, I was reflecting on it with a colleague, realizing that I need more organization. Entrepreneurship is a massive undertaking, and you must adapt a lot.

One of the participants clearly highlighted that investing in long-term organization limits the ability to adapt to uncertainty and the changing context of the modern world.

We don't make long-term plans because it would be a waste of time. Experience has taught us that if we plan for twelve months, after three weeks, so many things have happened in the market that we had to pivot or change direction.

Among the most commonly used digital productivity tools mentioned are Trello, Slack, D.Gantt, and Canva. However, in three cases, there is still a preference for using analogue tools. One of the reasons is the lack of preparedness of the team to manage work using this type of technology.

Not everyone internally is accustomed to it, especially if they are not from the technology field. In work teams, not everyone is ready to deal with this kind of burden.

Based on the previous results, we can conclude that although successful digital entrepreneurs are aware of the importance of using personal organization strategies, they lack the habits to apply them in practice. Among those who do use them, there is a diversity of opinions regarding the relevance of using digital tools versus analogue methods.

The third objective of the interview relates to the processes of implementing entrepreneurial action and collaborating with other agents. From the gathered data, we can observe that entrepreneurs do not have a clear understanding of what a digital strategy is. Among those who did (three cases), they indicated that they have not defined it, despite recognizing its importance.

We also noticed that the role of the Community Manager is not fully present because they consider it a lower priority, as indicated in eight cases. The most used communication tool is WhatsApp. Over half of the participants use surveys with their audience and the digital storytelling technique to communicate.

Regarding the digital tools used by entrepreneurs to disseminate digital content related to their idea, product, or company, we see that social media platforms (eleven cases) are the most used.

Finally, regarding the success keys mentioned by the interviewees to improve their digital entrepreneurship competence, they consider the following factors fundamental:

- Self-awareness of strengths and limitations (general agreement of all of them).
- Persistence (six cases).
- Continuous learning (five cases).
- Turning problems into learning opportunities (three cases).

To finish this section around the interviews, we would like to remark this idea expressed by one successful entrepreneur interviewed in the research: "For an entrepreneur, whether digital or not, it is essential to be consistent and not give up because it is a very long, long journey. We tend to be impatient and want to see new results".

5. DISCUSSION

In recent years, there has been a growing consensus that aligning life purpose with entrepreneurship (Vázquez, 2021) significantly increases the likelihood of success. Surprisingly, while this concept has gained traction, academic research confirming this correlation remains elusive, a finding that aligns with the results obtained in this study.

Interestingly, participants in this study offered somewhat superficial reflections on this seemingly critical phase. This phenomenon raises the intriguing proposition that, for entrepreneurs, iterative processes based on trial and error may be more effective than introspective processes rooted in pre-existing beliefs and values. This view of successful entrepreneurs would be in contradiction with previous research (Chatterjee et al., 2022). In addition, students indicate that they feel able to identify their opportunities, based on their own assessment. It is worth noting that in this case males showed a perception of feeling more capable than females. Some gender differences have also been found in previous studies on entrepreneurship (Adachi & Hisada, 2017; Camelo et al., 2016; Ruiz et al., 2021; Zesser et al., 2019) and even in studies based on our own model (Román-Garcia & González-Calatayud, 2022).

An alternative interpretation related to the distancing of successful entrepreneurs from planning arises from the recognition of the historical gap between academic theories of entrepreneurship and their practical application. Blank and Eckhardt (2023) have ventured to bridge this chasm by forging a connection between the Lean Startup methodological approach and the world of academic theories of entrepreneurship. Their argument is that both spheres can benefit from a closer relationship. The Lean Startup paradigm revolves around a continuous build-measure-learn cycle, enabling early validation of hypotheses and rapid refinement of products and services. It is now postulated as a scientific approach to accelerate the delivery of services and products to consumers. Different previous studies (Ghezzi & Cavallo, 2020; Chung, 2015; Harms, 2015; Tohanean & Weiss, 2019) attest to its effectiveness, not only in the broader business context,

but also in the realm of digital entrepreneurship. Surprisingly, there is even a specific guide for scientists and researchers aspiring to become entrepreneurial academics (Ruíz & García, 2023).

Other important issue in our model is the action plan and it is reflected both in interviews and questionnaires results. Thus, students highlight that they feel competent to plan, for example to design a digital entrepreneurship process. Therefore, we find ourselves with a key competence that students say they have mastered, although female students seem to feel more competent than their male peers (Román-García & González-Calatayud, 2022).

In the area of planning, successful digital entrepreneurs recognize the paramount importance of employing personal organizational strategies, which other authors call autonomy (Azqueta & Naval, 2019). However, when it comes to putting them into practice, they often find it difficult to establish effective habits. This contrasts with students' perception that they are able to put ideas into practice and solve problems. This difference in perception is probably due to a lack of actual implementation on the part of the students.

In addition, there is a divergence of opinion among these entrepreneurs as to the desirability of using digital tools versus adhering to traditional analogue methods. This also applies to students when asked about their use of tools for learning (Prendes-Espinosa et al., 2018). Interestingly, no existing study has delved into the intricate landscape of productivity strategies and effective organization in digital entrepreneurship, which sheds light on an attractive avenue for future research. It is worth noting that, by contrast, social media platforms abound with information on these very topics.

If we focus on the data provided by the students, it is curious to find that students show themselves to be capable when it comes to project management. In general, in all the competences included in this dimension, the results show a profile of students who are able to learn from experience, solve problems, organize themselves, motivate themselves and follow an ethical vision of their work (García-Tudela et al., 2022). If anything stands out about the students, it is the dimension of implementation and collaboration, especially considering that within this is the Communication and Collaboration competence. Standing out in this aspect seems to be a factor of possible future success when setting up a business, as Fernández (2023) states in his study. In neither of these two dimensions were significant differences found between men and women, with both feeling equally capable.

The entrepreneurs themselves highlight the virtue of persistence as a key factor underpinning their achievements, a sentiment that partially coincides with the findings of Santos et al. (2023). In contrast to our study, this body of work introduces the facets of self-promotion and social connectedness as fundamental pillars of entrepreneurial success. Notably, the entrepreneurs who participated in our research also highlighted other keys to their success, such as strong problem-solving skills, deep self-awareness and an unwavering commitment to lifelong learning. Several of these latter characteristics are aspects of the Em-Digital model (Prendes-Espinosa & García-Tudela, 2020) and the students surveyed seem to exhibit some of them, such as problem solving.

These characteristics of successful entrepreneurs related to those possessed by students are essential for improving students' digital entrepreneurship competence. Equipping students with sufficient skills to be able to become entrepreneurs is becoming essential in today's society (Čapienė & Ragauskaitė, 2017).

But these skills or training needs are not the same for all students. Most significantly, students show more or less entrepreneurial competence depending on the field of knowledge of the degree. Fostering entrepreneurship is a key factor for students to take the initiative (Scuotto & Morellato, 2013). This suggests that universities should make an effort to integrate this competence more transversally in these branches, providing resources to students and training teachers.

6. CONCLUSIONS

Entrepreneurship has become an essential aspect of today's society, but today it cannot be understood without its link to the digital world. Thus, digital entrepreneurship is the key point of this research and it must be also considered a basic skill in university students education. We started from the creation of a model called EmDigital, which is based on the EntreComp model (Bacigalupo et al., 2016) and the DigComp model (Vourikari et al., 2016), both from the European Commission. Based on this EmDigital model, the aim of our study was to find out the level of digital entrepreneurship competence of final-year undergraduate students. In addition, we wanted to know the opinion of successful digital entrepreneurs on the keys to success in a digital entrepreneurship project. The information collected from both profiles allows us to generate valuable results in terms of training related to digital entrepreneurship competence.

One of the main contributions to knowledge generated in this research is the EmDigital model itself, but also the research instruments (quantitative and qualitative) designed and validated on the basis of it. We have mainly used the European models EntreComp and DigComp, so this proposal is aligned with the European framework of competences for citizenship. Moreover, a substantial part of the interview design was inspired by the visual tool known as StrategicTree (Ogel & Castillo, 2012). This tool highlights the fundamental nature of pre-entrepreneurship introspection, which encompasses considerations of purpose, life experiences, personal beliefs, values and educational background. If we focus on the EmDigital model, this introspection task is observed in the dimension of identification of opportunities. This introspective task is seen as the basis for shaping innovative ideas, products and companies, so data collected can have a relevant educational and social impact.

Our model also reflects the process of generating and managing entrepreneurial projects, as the four competence areas can also be used as the four reference phases for the development and management of digital entrepreneurial projects. Thus, the next step is the need to generate a clear plan that serves as a route to follow, an aspect that successful digital entrepreneurs also highlight as an element to consider if we want to achieve success in our ideas, although without starting from the previous introspection mentioned above. And after the plan, we need skills to implement, to manage and to solve problems taking into account the other aspects as security, ethic or sustainability.

All these topics of interest have been remarked by successful digital entrepreneurs in the interviews, so have been able to identify the main elements and strategies used by them in their real projects. The other research objective was to analyze the self-perception of digital entrepreneurship competence of university students. Although some questions are open for future research, we found that in general, final year university students have an acceptable perception of digital entrepreneurship competence, with some differences

observed between the female and male population, as well as some differences between fields of knowledge. All these aspects will have to be studied in greater depth in the project on which we are currently working, with the approach of a more extensive and international data collection in order to be able to make comparative analyses.

7. FUNDING

This research has been funded by Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia" ("Competencias para el emprendimiento digital de los estudiantes universitarios", Reference: 20962/PI/18) and the Spanish Ministry of Science and Universities ("Marco de análisis del emprendimiento digital en las universidades", Reference PID2022-139332NB-I00) both under the direction of M.P. Prendes-Espinosa (University of Murcia, Spain).

8. ACKNOWLEDGEMENT

We would like to thank all the participants in this study. To the students who filled in the questionnaire and to the teachers who provided access to their classrooms. We would also like to thank all the entrepreneurs who agreed to be interviewed and the professionals of the "Instituto de Fomento de la Región de Murcia" for their collaboration. Moreover, to the "Servicio de Apoyo Estadístico" (University of Murcia) because its support for the quantitative analysis of data.

9. REFERENCES

Adachi, T., & Hisada, T. (2017). Gender differences in entrepreneurship and intrapreneurship: an empirical analysis. *Small Business Economics*, *48*, 447-486. https://doi.org/10.1007/S11187-016-9793-Y

Azqueta, A., & Naval, C. (2019). Entrepreneurship education: a proposal for human development. *Revista Española de Pedagogía, 77*(274), 517-534. https://doi.org/10.22550/REP77-3-2019-03

Bacigalupo, M. (2022). Competence frameworks as orienteering tools. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa*, (12), 20–33. https://doi.org/10.6018/riite.523261

Bacigalupo, M., Kampylis, P., Punie, Y., & Van den Brande, G. (2016). *EntreComp: The Entrepreneurship Competence Framework*. Joint Research Centre. https://doi.org/10.2791/593884

Basly, S., & Hammouda, A. (2020). Family Businesses and Digital Entrepreneurship Adoption: A Conceptual Model. *The Journal of Entrepreneurship, 29*(2), 326–364. https://doi.org/10.1177/0971355720930573

Blank, S., & Eckhardt, J. T. (2023). The Lean Startup as an Actionable Theory of Entrepreneurship. *Journal of Management, 0*(0), 1-23. https://doi.org/10.1177/01492063231168095

BOE (2020). Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. 30 de diciembre de 2020. *BOE, 340*, 122868-122953. https://www.boe.es/eli/es/lo/2020/12/29/3

Bogdanowicz, M. (2015). *Digital Entrepreneurship Barriers and Drivers. The need for a specific measurement framework*. Institute for Prospective Technological Studies, JRC Technical Report. https://doi.org/10.2791/3112

Camelo Ordaz, C., Diánez González, J. P., & Ruiz Navarro, J. (2016). The influence of gender on entrepreneurial intention: the mediating role of perceptual factors. *Business Research Quarterly, 19*(4), 261-277. https://doi.org/10.1016/j.brq.2016.03.001

Čapienė, A., & Ragauskaitė, A. (2017). Entrepreneurship education at university: innovative models and current trends. *Research for rural development, 2*, 284-291. https://doi.org/10.22616/RRD.23.2017.080.

Carreón, J. (2014). Emprendedurismo migrante y comerciante. *Tlatemoani*, (15), 1-30

Chatterjee, S., Chaudhuri, R., Vrontis, D., & Basile, G. (2022). Digital transformation and entrepreneurship process in SMEs of India: a moderating role of adoption of AI-CRM capability and strategic planning. *Journal of Strategy and Management*, *15*(3), 416-433. https://doi.org/10.1108/JSMA-02-2021-0049

Chung, J. (2015). Lean startup for academic entrepreneurship: Building a scalable startup launchpad. In A. Szopa, W. Karwowski and D. Barbe (Coord.), *Competitive strategies for academic entrepreneurship: Commercialization of research-based products* (pp. 171-187). IGI Global. https://doi.org/10.4018/978-1-4666-8487-4.ch008

Cruz García, L. (2015). Especificación de un modelo de emprendimiento electrónico. *Tlamati, 6*(1), 72-83. http://tlamati.ua/gro.mx/t61/t6110.pdf

Cruz García, L. (2016). Emprendimiento digital: estudio de caso con universitarios de comunicación, UAEM UAP Huehuetoca. *Revista de Ciencias Sociales*, *29*, 34-45. https://revistas.upr.edu/index.php/rcs/article/view/7395/6021.

European Commission. (2006). *Competencias clave para el aprendizaje permanente*. Recomendación 2006/962/CE del Parlamento Europeo y del Consejo, de 18 de diciembre de 2006, sobre las competencias clave para el aprendizaje permanente [Diario Oficial 394 de 30.12.2006] https://bit.ly/3tUNIO2

European Commission. (2018). *EntreComp into Action: get inspired, make it happen.* http://ec.europa.eu/social/main.jsp?catId=1317&langId=en

Fernández, M. V. (2023).Competencias digitales clave en el emprendimiento juvenil: una revisión sistemática de los últimos 6 años. *RiiTE, Revista Interuniversitaria de Investigación en Tecnología Educativa, 14*, 28-44. https://doi.org/10.6018/riite.565401

Fossati, P., Chiappetta, C.J., Ratten, V., Mederios, G., Borchardt, M., Sperandio, G., & Eberle, L. (2023). What do (should) we know to leverage students' employability and entrepreneurship? A systematic guide to researchers and managers. *The International Journal of Management Education, 21*(2), 1-15. https://doi.org/10.1016/j.ijme.2023.100788

García-Tudela, P. A., González-Calatayud, V., & Montiel-Ruiz, F. J. (2022). Aprendizaje autodirigido y emprendimiento: una aproximación desde la realidad universitaria. *Educar, 58*(2), 405-425. https://doi.org/10.5565/rev/educar.1518

García-Tudela, P.A., González-Calatayud, V., & Prendes-Espinosa, M. P. (2020). El emprendimiento digital como competencia para la formación de los estudiantes universitarios. En G. Gómez, M.R. Navas-Parejo, C. Rodríguez y J.C. De la Cruz (Eds.), *Teoría y práctica en investigación educativa: una perspectiva internacional* (pp. 2207-2216). Dykinson. https://bit.ly/3ezel2J

GEM (2020). *Global entrepreneurship monitor. Informe GEM España 2020*. https://bit.ly/3N96J6k

Ghezzi, A., & Cavallo, A. (2020). Agile business model innovation in digital entrepreneurship: Lean startup approaches. *Journal of Business Research, 110*, 519-537. https://doi.org/10.1016/j.jbusres.2018.06.013

Giones, F., & Brem, A. (2017). Digital Technology Entrepreneurship: A Definition and Research Agenda. *Technology Innovation Management Review*, *7*(5), 44–51.

González-Calatayud, V., Román García, M., & Prendes Espinosa, M. P. (2018). Formación en competencias digitales para estudiantes universitarios basada en el modelo DigComp. *Edutec. Revista Electrónica de Tecnología Educativa, 65*, 1-15. https://doi.org/10.21556/edutec.2018.65.1119

Gutiérrez-Mora, X. A., Castro-López, G. A., Viteri-Medina, A. A., & Parra-Bustamante, M. F. (2021). El perfil psicológico del emprendedor y su proyección ante tiempos de crisis. *FIPCAEC, 6*(3), 76-90. https://doi.org/10.23857/fipcaec.v6i1.429

Harms, R. (2015). Self-regulated learning, team learning and project performance in entrepreneurship education: Learning

in a lean startup environment. *Technological Forecasting and Social Change, 100*, 21-28. https://doi.org/10.1016/j.techfore.2015.02.007

Jardim, J. (2021). Entrepreneurial Skills to Be Successful in the Global and DigitalWorld: Proposal for a Frame of Reference for Entrepreneurial Education. *Education Sciences*, *11*(7), 1-13. https://doi.org/10.3390/educsci11070356

Kollman, T. (2009). E-entrepreneurship: the principles of founding electronic ventures. En I. Lee (Ed.), *Selected readings on information technology and business systems management* (pp. 204-225). IGI Global. https://dx.doi.org/10.4018/978-1-59904-901-4.ch008

Lorenzo, O. (2012). La empresa en Red: emprendimiento+innovación+Tecnología. *Boletín de estudios económicos, 67*(205), 133-144. http://bit.ly/2qUl1Dt

Mababu, R. (2017). La transformación digital y el emprendimiento de los jóvenes en Iberoamérica. *Revista Internacional y Comparada de Relaciones Laborales y Derecho del Empleo, 5*(2), 111-128.

Maxwell, J. A., & Miller, B. A. (2008). Categorizing and connecting strategies in qualitative data analysis. In S. N. Hesse-Biber y P. Leavy (Eds.), *Handbook of emergent method*s (pp. 461–477). The Guilford Press.

McCallum, E., Weich, R., McMullan, L., & Price, A. (2018). *Get inspired make happen intoAction EntreComp. A user guide to the European Entrepreneurship Competence Framework*. Publication Office of the European Union. https://doi.org/10.24320/redie.2018.20.1.1347

Nambisan, S. (2017). Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship. *Entrepreneurship Theory and Practice*, *41*(6), 1029–1055. https://doi.org/10.1111/etap.12254

Ogel, L., & Castillo, A. (2012). *Árbol estratégico. Diseñar empresas fundamentadas en valores.* Fondo Social Europeo y EOI.

Omar, F. I., Othman, N. A., & Hassan, N. A. (2019). Digital inclusion of ICT and its implication among entrepreneurs of small and médium enterprises. *International Journal of Engineering and Advanced Technology, 8*(5), 747-752. http://dx.doi.org/10.35940/ijeat.E1106.0585C19

Pérez, M., Carreón, J., Quintero, M. L., Bucio, C., García, C., & Aguilar, J. A. (2016). La agenda institucionalista y la gestión del conocimiento: especificación de un modelo de emprendimiento innovador. *KAIROS Revista de Temas Sociales,* (38), 29-37.

Prendes-Espinosa, M. P. (coord.) (2022). *Emprendimiento digital*. Tirant Lo Blanch.

Prendes-Espinosa, M. P., & García-Tudela, P. A. (2020). Modelo EmDigital: áreas e indicadores de la competencia de emprendimiento digital. En E. Archundia, M. A. León, y C. Cerón (Eds.), *Redes de aprendizaje digital en nodos colaborativos* (pp. 361-372). Benemérita Universidad Autónoma de Puebla. https://hdl.handle.net/20.500.12371/9982

Prendes-Espinosa, M. P., Solano-Fernández, I. M., Serrano-Sánchez, J. L., González-Calatayud, V., & Román-García, M. del M. (2018). Entornos Personales de Aprendizaje para la comprensión y desarrollo de la Competencia Digital: análisis de los estudiantes universitarios en España. *Educatio Siglo XXI*, *36*(2 Jul-Oct), 115–134. https://doi.org/10.6018/j/333081

Prendes-Espinosa, P., Solano-Fernández, I. M., García-Tudela, P. A. (2021). EmDigital to Promote Digital Entrepreneurship: The Relation with Open Innovation. *Journal of Open Innovation: Technology, Market and Complexity, 7*(1), 1-14. https://doi.org/10.3390/joitmc7010063

Ratten, V., & Usmanji, P. (2021). Entrepreneurship education: time for a change in research direction? *The International Journal of Management Education*, *19*, 1-8. https://doi.org/10.1016/j.ijme.2020.100367

Román-García, M. M., & González-Calatayud, V. (2022). La competencia de emprendimiento digital en función del género: el proyecto EmDigital. *Hachetetepé. Revista científica de Educación y Comunicación, 24,* 1-13. https://doi.org/10.25267/Hachetetepe.2022.i24.1205

Ruiz Navarro, J., Camelo Ordaz, C., & Coduras Martínez, A. (2021). Mujer y desafío emprendedor en España: características y determinantes. *Economía industrial*, 383, 13-22. http://tinyurl.com/yk9pte4s

Ruiz, A. (2012). Aprende a emprender: Universidad y emprendimiento. *Revista de Estudios de Juventud*, *99*, 69-87. http://tinyurl.com/yn52wtrb

Santos, S. C., Caetano, A., & Brochado, A. (2023). Why am I so successful? Self-presentation and deliberate attributions of success in entrepreneurship. *Journal of Business Venturing Insights, 19*, 1-14. https://doi.org/10.1016/j.jbvi.2023.e00396

Satalkina, L., & Steiner, G. (2020). Digital Entrepreneurship and its Role in Innovation Systems: A Systematic Literature Review as a Basis for Future Research Avenues for Sustainable Transitions. *Sustainability, 12*, 2764, 1-27. https://doi.org/10.3390/su12072764

Scuotto, V., & Morellato, M. (2013). Entrepreneurial Knowledge and Digital Competence: Keys for a Success of Student Entrepreneurship. *Journal of Knowledge Economy, 4*, 292-303. https://doi.org/10.1007/s13132-013-0155-6

Secundo, G., Rippa, P., & Cerchione, R. (2020). Digital Academic Entrepreneurship: A structured literature review and avenue for a research agenda. *Technological Forecasting & Social Change*, *157*, 1-17. https://doi.org/10.1016/j.techfore.2020.120118

Standing, C., & Mattsson, J. (2018). "Fake it until you make it": business model conceptualization in digital entrepreneurship. *Journal of Strategic Marketing*, *26*(5), 385-399.

Tohanean, D., & Weiss, P. (2019). Digital entrepreneurship and green business model innovation: Lean startup approaches. *Quality - Access to Success, 20*(S2), 630-634.

Vázquez, M. (2021). *El arte de emprender: ideas para diseñar tu negocio y tu vida*. Kindle.

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 47-73 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20896

Assessing the Impact of Microsoft Copilot and ChatGPT on EFL Learners' Interactional Metadiscourse in Argumentative Writing

Evaluación del impacto de Microsoft Copilot y ChatGPT en el discurso metadiscursivo interactivo de los estudiantes de inglés como lengua extranjera en la escritura argumentativa

RECEIVED 27/11/2024 ACCEPTED 26/02/2025 PUBLISHED 01/06/2025

Rajab Esfandiari
ORCID 0000-0002-2305-762X
Imam Khomeini International University, Iran
esfandiari@hum.ikiu.ac.ir

Omid Allaf-Akbary
ORCID 0000-0002-9019-9273
University of Mohaghegh Ardabili, Iran
oallafakbary@gmail.com

ABSTRACT

The rise of artificial intelligence (AI) chatbots has significantly transformed the educational landscape, offering numerous opportunities for innovation and change. The current study assessed the comparative effects of ChatGPT-based instruction and Microsoft Copilot in helping Iranian English-as-a-foreign language (EFL) learners identify and realize interactional metadiscourse markers (IMMs) in argumentative writing and exploring their attitudes towards these two chatbots. Grounded in the theoretical framework of IM, this study followed a convergent parallel design. The study involved 90 male and female language learners randomly assigned to three groups: ChatGPT-based group ($n = 30$), Microsoft Copilot group ($n = 30$), and a control group ($n = 30$). The experimental groups were provided with 10 prompts per session for the implementation of IMMs, resulting in 60 prompts across six sessions. Instruction included initial training on using the respective AI tools, followed by practice sessions focusing on identifying and using IMMs in writing. The control group received conventional instruction, which involved identifying IMMs in reading passages with the guidance of the instructor. Interview questions were designed to elicit perspectives from learners on their experiences with ChatGPT and Microsoft Copilot. The responses from the interview data concerning learners' perceptions were analyzed through thematic analysis. Results showed that the Microsoft Copilot group demonstrated superior performance compared to the other two groups in identifying IMMs in the posttest. However, a one-way analysis of covariance (ANCOVA) showed that the difference between the ChatGPT-based group and the control group was not statistically significant. Additionally, responses to semi-structured interviews indicated that all learners had a positive perception of Microsoft Copilot for employing IMMs in argumentative writing. This study contributes to the field by providing empirical evidence on the effectiveness of specific AI-driven chatbots in enhancing critical writing skills, specifically through the lens of IMMs.

KEYWORDS AI-enhanced learning; Metadiscourse realization; Semi-structured interview.

RESUMEN

El auge de los chatbots de inteligencia artificial (IA) ha transformado significativamente el panorama educativo, ofreciendo numerosas oportunidades para la innovación y el cambio. El presente estudio evaluó los efectos comparativos de la instrucción basada en ChatGPT y Microsoft Copilot en ayudar a los aprendices iraníes de inglés como lengua extranjera (EFL) a identificar y realizar marcadores de metadiscurso interactivo (IM) en la escritura argumentativa y a explorar sus actitudes hacia estos dos chatbots. Basado en el marco teórico de IM, este estudio siguió un diseño paralelo convergente. El estudio involucró a 90 aprendices de idiomas, tanto hombres como mujeres, asignados aleatoriamente a tres grupos: grupo basado en ChatGPT (n = 30), grupo de Microsoft Copilot (n = 30) y un grupo de control (n = 30). A los grupos experimentales se les proporcionaron 10 consignas por sesión para la implementación de marcadores IM, resultando en 60 consignas a lo largo de seis sesiones. La instrucción incluyó un entrenamiento inicial en el uso de las herramientas de IA respectivas, seguido de sesiones prácticas enfocadas en identificar y usar marcadores IM en la escritura. El grupo de control recibió instrucción convencional, que implicaba identificar marcadores IM en pasajes de lectura con la guía del instructor. Se diseñaron preguntas de entrevista para obtener perspectivas de los aprendices sobre sus experiencias con ChatGPT y Microsoft Copilot. Las respuestas de los datos de la entrevista sobre las percepciones de los aprendices fueron analizadas mediante análisis temático. Los resultados mostraron que el grupo de Microsoft Copilot demostró un rendimiento superior en comparación con los otros dos grupos en la identificación de marcadores IM en el post-test. Sin embargo, un análisis de covarianza (ANCOVA) mostró que la diferencia entre el grupo basado en ChatGPT y el grupo de control no fue estadísticamente significativa. Además, las respuestas a las entrevistas semiestructuradas indicaron que todos los aprendices tenían una percepción positiva de Microsoft Copilot para emplear marcadores IM en la escritura argumentativa. Este estudio contribuye al campo proporcionando evidencia empírica sobre la efectividad de chatbots impulsados por IA específicos en la mejora de habilidades críticas de escritura, específicamente a través de la lente de IM.

PALABRAS CLAVE Aprendizaje mejorado con IA; Realización del metadiscurso; Entrevista semiestructurada.

1. INTRODUCTION

The integration of AI (artificial intelligence) into educational systems has brought about significant changes, enhancing both teaching and learning processes. AI-powered tools offer personalized learning experiences, adapting to the unique needs of each student. These tools analyze students' performance and provide tailored feedback, helping them to overcome specific challenges and improve their skills (LaRue Keeley, 2024). Recent studies (e. g., Lo et al., 2024) have shown that AI-driven educational technologies can improve learning outcomes, increase accessibility, and foster a more inclusive learning environment. By leveraging machine learning algorithms, educational platforms can predict student performance, identify at-risk students, and provide timely interventions to support their academic progress (Sadiku et al., 2022). Thus, the impact of AI on education is profound, transforming traditional educational models and paving the way for innovative pedagogical approaches (Holmes et al., 2022).

This study explores the uncharted territory of how AI-assisted chatbots, specifically ChatGPT and Microsoft Copilot, enhance interational metadiscourse markers (IMMs) in EFL learners' argumentative writing. This research goes beyond theoretical discussions by providing empirical evidence of the effectiveness of these AI tools in improving critical writing skills. By comparing ChatGPT and Microsoft Copilot within an educational context, this research showcases their advanced natural language processing capabilities,

which allow for more nuanced and context-aware interactions than traditional AI tools. This ability to provide personalized feedback is particularly beneficial for fostering IMMs, which is often lacking in more rigid, rule-based systems. The real-time, adaptive feedback they offer can lead to more engaging and effective learning experiences, thus improving educational outcomes. The continuous updates and improvements in ChatGPT and Microsoft Copilot ensure they remain at the forefront of technological advancements, providing cutting-edge solutions that evolve with educational demands. This dynamic nature sets them apart from other AI tools that may not receive regular updates or enhancements, thereby maintaining their relevance and effectiveness in educational contexts.

Lo et al. (2024) ssytematiclly reviewed the appliaction of AI (e.g., ChatGPT) in the English as a second/foreign language education in 70 empirical research studies and found that the majority of the studies (n = 29) had focused on the students' use of AI tools in writing, including, but not limited to, the generation of ideas in writing, writing organization and structure, and spelling and grammar in writing. Lo et al.'s systematic review highlights the potential benefits of using AI tools to enhance different facets of writing skills. However, there is a growing need to extend this line of research to explore the impact of AI on interactional metadiscourse (IM) in argumentative writing. Investigating how AI tools like ChatGPT and Microsoft Copilot can assist learners in effectively using IMMs in their writing provide a deeper understanding of the role of AI in facilitating more nuanced and coherent written communication. This study may contribute to the broader field of language learning and offer practical insights for educators and learners seeking to leverage AI for advanced writing skills development. Therefore, more research studies are needed to help language teachers to gain a better understanding of how these AI-powered chatbots can be safely implemented in language classes to facilitate the teaching/learning process. Against this background, the current research examined how the application of AI chatbots contributes to the employment of IM in argumentative writing abilities of English-as-a-foreign language (EFL) learners.

This research study is innovative in several keyways. Firstly, it provides a comparative analysis of two leading AI models in the context of language teaching and learning, offering valuable insights into their relative strengths and weaknesses. Secondly, the study employs a detailed examination of transformer model architectures and training data updates, shedding light on the underlying causes of performance differences. Finally, this research contributes to the ongoing discourse on AI-assisted education by highlighting the potential of AI tools to enhance specific aspects of academic writing, thereby informing future developments in AI-driven educational technologies. Overall, the combination of advanced language processing, personalized feedback, and continuous improvement make ChatGPT and Microsoft Copilot superior options for enhancing EFL learners' IMMs in argumentative writing. The study provides valuable insights for educators and policymakers aiming to implement effective AI-assisted strategies in their curricula, potentially leading to more effective teaching methods and improved learning outcomes.

The significance of of using AI tools in the realization of IMMs for Iranian EFL learners lies in their capability to enhance their argumentative writing skills. Iranian language learners follow discourse and linguistic patterns in their argumentation, which may differ from those native-English writers use (Pearson & Abdollahzadeh, 2023). Mastering IMMs helps bridge such discoursal differences, making their writing more effective and culturally appropriate. It also improves critical thinking and argumentation skills, which are crucial

for presenting ideas clearly, anticipating reader objections, and providing appropriate rebuttals. Lastly, in the globalized world, persuasive English writing is a valuable skill, enabling Iranian EFL learners to engage effectively in academic and professional settings. Highlighting these points underscores the importance of this study for English language education in Iran. Therefore, the following research questions were raised in order to accomplish the purposes.

1. Are there any significant differences between ChatGPT-based instruction, Microsoft Copilot, and conventional instruction in employing IMMs among Iranian advanced EFL learners while producing argumentative writing?
2. How do the learners in experimental groups perceive ChatGPT-based instruction and Microsoft Copilot as feedback and learning tools while using IMMs in their argumentative writing skills?

2. LITERATURE REVIEW

In this section, firstly, we explain the theory we have used to base this study on and disucss how it is related to the present study. Next, Hylands' (2019) model of interactional metadiscourse, as used in the study, is introduced and elaborated on, followed by summaries of the fidnings of the empirical studies on metadiscourse, and the critical evaluation of these studies. Lastly, the role of AI tools in language education, most specifically ChatGPT and Microsoft Copilot, is explained, and the research studies using AI in language education are summarised, and their limitations are discussed.

2.1. Theoretical framework of the study

Feedback theory in language learning, which emphasizes the importance of specific, timely, constructive, and informative feedback, significantly enhances learners' linguistic abilities by helping them identify their progress and areas needing improvement. Studies like those by Nassaji and Kartchava (2024) and Moser (2020) have underscored the role of feedback in fostering engagement, self-regulation, and performance in learners. This process promotes a growth mindset where challenges are viewed as opportunities for development. Meanwhile, GPT-4 and Microsoft Copilot as writing assistants offer a practical application of feedback theory by providing immediate, precise, and constructive feedback to writers. GPT-4 and Microsoft Copilot aid in idea generation, content creation, editing, and proofreading, thereby not only enhancing efficiency and creativity but also aligning with the principles of effective feedback (Van, 2023). For instance, when GPT-4 and Microsoft Copilot suggest improvements or identifies errors, they act as forms of corrective feedback that writers can utilize to refine their work, as a human tutor or an editor does. This immediate feedback loop supports writers in making real-time adjustments, fostering a continuous learning process that mirrors the benefits highlighted in feedback theory (Mackey, 2020). By integrating these feedback mechanisms, GPT-4 and Microsoft Copilot help maintain consistency, accuracy, and originality in writing while addressing quality control and ethical considerations. In essence, the synergy between feedback theory and GPT-4 and Microsoft Copilot as writing assistants demonstrate how advanced AI tools can operationalize theoretical principles of feedback, providing writers with support and guidance needed to enhance their skills and produce high-quality work (Bautista, 2024). This combined approach underscores

the transformative potential of leveraging feedback and AI technology to foster improvement in both language learning and writing processes.

The present research was grounded in feedback theory process (Winstone & Carless (2020) and Barrot's (2023) ChatGPT speculation, with the objective of employing GPT-4 and Microsoft Copilot as tools for providing formative feedback that incorporates both self-assessment and peer-assessment. The conceptual framework for utilizing GPT-4 and Microsoft Copilot to improve formative feedback can be examined through two prominent theories. The first is the dialogic feedback model articulated by Winstone and Carless (2020), who contend that feedback characterized by interactive dialogue allows students to clarify their expectations, obtain essential information and guidance, and enhance their learning outcomes. GPT-4 and Microsoft Copilot facilitate this interactive process by responding to user inquiries regarding various aspects of writing, providing suggestions as needed, and acting as an accessible support tool. Additionally, it identifies and rectifies errors, thereby enriching the interactive exchange. The second theory is Barrot's (2023) view of GPT-4 and Microsoft Copilot as reliable writing assistants that offer prompt, context-sensitive, and personalized feedback to students at various stages of their writing endeavors.

The current study followed a comparative study of AI-based assessments (ChatGPT and Microsoft Copilot) as feedback tools to help learners realize IMMs and assess their impact on learners' argumentative writing skills. For this reasosn, we now turn to metadiscourse to explain its princples and summarise the fidnings of previous studies in this area.

2.2. Interactional metadiscourse

Metadiscourse serves an essential function in achieving persuasion in all forms of human communication. Despite the seemingly impersonal nature of academic writing, establishing a suitable connection between the author and the audience is essential for effectively conveying a credible argument (Qiu et al., 2024). In order to achieve this goal, writers must utilize rhetorical strategies that are persuasive to their audience. Metadiscourse is a term frequently utilized in discourse analysis, denoting features of a text that pertain to its structure or substance (Ädel, 2023). Metadiscourse involves the linguistic strategies utilized by writers or speakers to influence how a recipient understands a written or spoken work. Following Hyland's (2019) interpersonal metadiscourse, the current research shed light on the IMMs (table 1). Regarding argumentative writing, IMMs are crucial to support the writer's ideas and aid the reader's understanding (Abdelrahim & Abdelrahim, 2020).

TABLE 1. Interactional metadiscourse in Hyland's model (Hyland, 2019, p. 58)

Interactional metadiscourse	Definition	Examples
Hedges	refrain from commitment	perhaps; about
Boosters	emphasize assurance	certainly; clearly
Attitude markers	conveying the writer's stance	unfortunately; I agree
Self-mentions	author mention	I; my
Engagement markers	intentionally foster a connection with the reader	consider; note

IMMs pertain to the methods by which writers engage in interaction through the intrusion and commentary on their own messages (Hyland, 2019). The primary objective for writers is to clarify their perspectives and actively involve readers by inviting them to react to the developing narrative. This process reflects the writer's distinctive persona recognized within the community, encompassing the manner in which they articulate judgments and explicitly align themselves with their audience. IMMs serve a fundamentally evaluative and participatory role, demonstrating solidarity, anticipating counterarguments, and engaging in a hypothetical dialogue with others (Gordon, 2024). It illustrates the degree to which the writer endeavors to collaboratively construct the text alongside the readers.

To enhance their understanding of metadiscourse, scholars have undertaken investigations in a variety of scholarly genres including research articles (RAs) (Hyland & Jiang, 2022; Li & Xu, 2020; Pearson & Abdollahzadeh, 2023), concluding that metadiscourse has been essential in organizing discourse, involving the audience. The findings of these studies showed that in RAs writers tend to engage with their audience in an impersonal manner rather than a personal tone. This preference for impersonal metadiscourse can be attributed to the characteristics of research article writing, where the primary function of metadiscourse is to facilitate the reading experience by signaling the structure of the discourse and elucidating the relationships and meanings of the propositions presented. Other studies adopted an experimental approach such as the incorporation of learning-oriented language assessment into metadiscourse (Esfandiari & Allaf-Akbary, 2024a), explicit and implicit teaching of IMs (El-Dakhs, et al., 2022), and metadiscourse learning through AI-powered chatbots (Esfandiari & Allaf-Akbary, 2024b), stating that ChatGPT-based instruction was useful in learning and employing IMMs and learners had positive viewpoints towards ChatGPT-based instruction.

Ho and Li (2018) conducted research to analyze 181 argumentative essays written by first-year university students in a timed writing task using the interpersonal model of metadiscourse. The findings showed that low-rated essays had difficulties using metadiscourse to construct convincing arguments compared to high-rated essays. The study suggested that teaching and learning metadiscourse should be implemented in secondary and early tertiary education to help learners effectively use metadiscourse in academic writing.

Esfandiari and Khatibi (2022) analyzed 240 English academic RAs by American and Persian scholars to study IM. Using Hyland's (2019) interpersonal model, Esfandiari and Khatibi examined common features and different purposes in abstracts, introductions, and conclusions across American and Persian international and national corpora. Similarities were found in metadiscourse preferences between American and Persian academics, but differences in marker utilization featured as well. The results suggested that cross-cultural influences play a significant role in the frequency and utilization of interactional resources.

In their study, Rostami Aboo saeedi et al. (2023) conducted interviews with Iranian EFL graduate students to examine their perspectives on using IMMs in their theses. A total of 20 participants, including 15 females and 5 males, were interviewed. The interviews focused on topics such as general concepts, audience, language support, organization, and attitude. The analysis of the interviews revealed five main themes: general, audience-related, language support, organization, and attitude. Participants expressed a preference for research topics that are engaging and practical. Using attitude markers, engagement markers, and hedges effectively aids expressing viewpoints in theses, as stated by interviewees. They also discussed their challenges with writing coherence and cohesion in their theses, highlighting the importance of using markers to convey their viewpoints effectively.

In another more recent study, Qui et al. (2024) investigated IM, including hedges, boosters, attitude markers, and self-mentions in a corpus of L1-English expert and L1-Chinese student writing in Agricultural Science. Both groups showed differences in the use of these metadiscourse features, with L1-English experts using more hedges and L2 students using more boosters and attitude markers. Despite some similarities in functional subtypes, L1-English experts excelled in stating a goal or purpose in self-mentions. The analysis revealed discipline-related metadiscourse challenges for students. The study suggested the need for improved coding methods and highlights the importance of teaching metadiscourse to discipline-specific writers using relevant corpora.

Despite the significant advancements in understanding metadiscourse within academic contexts, as reviewed above, which underscore the role of impersonal engagement in structuring academic discourse, critical gaps remain that necessitate further investigation, particularly in the context of AI-assisted learning tools. The above-reviewed studies emphasize the structural importance of metadiscourse but have not explored the potential of AI tools to transform these dynamics by providing interactive and personalized engagement. While experimental approaches by Esfandiari and Allaf-Akbary (2024b) have demonstrated the effectiveness of AI in teaching metadiscourse, their study did not compare the benefits of different AI-chatbots. Ho and Li (2018) highlighted the struggle which students face in using metadiscourse effectively to construct convincing arguments, pointing to the need for early education in metadiscourse that the current research seeks to address by evaluating how different AI tools can enhance argumentative writing skills through consistent guidance. By focusing on AI-assisted learning environments, this study addresses the limitations of previous studies that often isolate metadiscourse from its broader communicative context.

Examining the specific impact of two different AI tools on EFL learners' interactional metadiscourse in argumentative writing, our study aims to provide new insights and practical applications, enhancing metadiscourse learning and usage among EFL learners. Therefore, in the next section, we will explain AI tools and present a critical discussion of the studies in this area.

2.3. AI-powered Chatbots in language education

As new tools continue to emerge with the advancement of technology, one of the most cutting-edge and contemporary technologies is AI. AI encompasses a class of systems which can be mechanized by utilizing technologies such as intelligent retrieval and statistical procedures (Lee et al., 2023). These systems have the capability to assist learners in performing tasks more quickly and accurately than they could on their own (Shneiderman, 2022). The advancement of AI-powered technologies has simplified the process for learners to receive feedback on their written work (Marzuki et al., 2023). The employment of AI chatbots such as ChatGPT to enhance feedback in writing courses is an emerging field that warrants additional research (Barrot, 2023). ChatGPT's role as a learning aid can be linked to constructivism. The constructivist learning theory asserts that individuals actively build their understanding by interacting with new information, which they then assimilate with their prior knowledge (McCourt, 2023). Allagui (2023) suggests that ChatGPT has the ability to aid learners in their writing by offering suitable guidance on content and structure while they are composing their work. Learners have the opportunity to obtain personalized feedback that is appropriate for their specific requirements. Integrating AI literacy across academic disciplines helps

students connect AI knowledge with their field-specific learning, fostering a comprehensive understanding of AI's impact and developing critical evaluation skills (Kim et al., 2022). To bridge the theoretical gap, we have explained how AI chatbots' advanced natural language processing capabilities can enhance IMMs. For instance, AI tools can analyze and generate texts with nuanced context awareness, offering learners immediate and contextually appropriate feedback on their use of IMMs. This connection is supported by studies that show how technology-mediated feedback can improve learners' writing skills by providing targeted guidance and promoting self-regulated learning (Xue, 2024).

Some studies have been conducted to highlight the significance of AI in learning lexical items and language structure (Wang & Guo, 2023), in giving feedback and invaluable data (Dai et al., 2023; Rudolph et al., 2023), promoting learners' motivation (Ali et al., 2023), in creating fine texts (Gao et al., 2023), in developing learners' writing skill (Mahapatra, 2021; Yan, 2023), thereby showing learners' positive attitude towards AI, in personalized language learning (Amirjalili et al., 2024; Esfandiari & Allaf-Akbary, 2024c), in improving EFL speaking practice (Clark, 2024), demonstrating a positive influence on learners' speaking abilities, encompassing accuracy and fluency. Huang et al. (2023), for example, carried out a comparative analysis to evaluate the academic achievement and learners' involvement in an AI-enhanced classroom in contrast to a traditional classroom without AI technology. The outcomes demonstrated that learners in the AI-driven classroom revealed better learning results and increased levels of engagement, compared with their counterparts in the non-AI classroom. Huang and Zou (2024) examined the roles of enjoyment and readiness to engage in communication with AI. The results showed a positive and useful impact of AI in enhancing learners' communication.

Several limitations emerge from the studies conducted on AI writing tools and their impact on EFL learners. For instance, Zhao (2022) critiques the focus of digital writing tools on revising and editing, highlighting a scarcity of support during the actual writing process, which may limit the development of holistic writing skills in learners. Moreover, Marzuki et al. (2023) highlight the positive perceptions of EFL teachers towards various AI writing tools, but they rely on qualitative data from a small sample, which may not fully capture broader trends and efficacy across diverse educational settings. Similarly, Escalante et al. (2023) compare AI-generated feedback with human tutors, finding no significant difference in outcomes, suggesting that a blended approach might be most effective. However, the study does not address the nuances of integrating such feedback seamlessly into traditional teaching methods.

Several other studies, as reviewed in this section, have focused on technology acceptance and its impact on learners. Wu et al. (2024), for example, employ the technology acceptance model and reveal that perceived ease of use significantly influences learners' attitudes towards AI. However, the study indicates that perceived usefulness did not have a significant predictive value, suggesting potential gaps in aligning AI tool functionalities with learners' specific needs. Furthermore, Huang et al. (2024) underscore the role of generative AI acceptance and teachers' enthusiasm in enhancing learners' well-being and self-efficacy. Nevertheless, the study does not address practical implementation strategies to sustain these positive effects. Additionally, Kim et al. (2024) identify both benefits and challenges in using a ChatGPT-4 integrated system but do not extensively explore long-term impacts on learners' writing performance and emotional engagement. Boudouaia et al. (2024) demonstrate the effectiveness of ChatGPT-4 in improving EFL writing skills,

but the controlled experimental design may not reflect real-world classroom dynamics and potential challenges in AI adoption. Lastly, Mahapatra (2024) reports positive impacts of ChatGPT on ESL students' writing skills but emphasizes the need for proper training, highlighting a potential gap in resource availability and accessibility for effective AI tool utilization.

3. METHOD

3.1. Population and sample

The current study consisted of 90 advanced male and female EFL learners aged between 25 and 28. Their first language was Turkish. These learners were registered at the academic center for education, culture, and research (ACECR) in Ardabil, Iran. The participants were chosen through convenience sampling from a larger pool of 142 learners, built on their performance in Michigan Test of English Language Proficiency (MTELP) (Lim, 2011). Advanced learners are more capable of engaging deeply with AI tools like ChatGPT and Microsoft Copilot, which are designed for high-level language processing and feedback. This proficiency level ensures that the impact of these tools can be more clearly observed and measured. None of the participants had already utilized ChatGPT and Microsoft Copilot to enhance their proficiency in writing. Other inclusion criteria involved extensive years of learning experience, and a strong background in academic writing, evidenced by coursework and writing samples. We also considered their professional writing experience, self-assessment of writing skills, recommendations from language instructors, and previous research participation. These comprehensive criteria ensure the participants possess a high level of language competence, enhancing the validity and reliability of our study's findings in terms of the impact of AI platforms on their interactional metadiscourse in writing.

Before the treatment sessions began, the participants were briefed on the objectives of the research and what was expected of them. In order to guarantee the participants' proficiency level, all participants were required to take MTELP. Learners scoring above 70 percent were categorized as advanced level language learners (Phakiti, 2003). Ultimately, the participants were randomly assigned into three different groups: 30 participants in the ChatGPT group, 30 in the Microsoft Copilot group, and 30 in the control group. All groups received printed materials on introductory concepts of IMMs and their subcategories. The experimental groups shared 15 personal computers (one for every two participants) connected to the internet. They received 10 prompts per session related to IMMs, using ChatGPT and Microsoft Copilot over six sessions, totaling 60 prompts, with a posttest in the final session. Prompts included identifying verbs for 'Hedges', providing examples of 'Hedges' and 'Boosters', and writing paragraphs including IMMs. Both experimental groups used the same prompts and received responses from their respective AI systems.

Prior to the grouping procedures and instructional courses, all participants underwent a metadiscourse pretest. This pretest involved analyzing four samples of TOEFL argumentative writing paragraphs to identify and underline IMMs. Participants who mistakenly underlined words with propositional meaning were randomly assigned to a study group, while those who correctly identified IMMs were not included in the further analysis. Before engaging in the study, the participants provided their consent to confirm that their participation was voluntary. At the outset of the treatment session, the participants received detailed information

regarding the specific aims of the research. Treatment intervention lasted for nine sessions (three sessions per week). The first two sessions were to teach how to use AI-powered chatbots (following Mathur & Mahapatra, 2022, advocating the provision of training for students prior to the implementation of a digital tool in educational settings) and introductions of metadiscourse. Each session took two hours.

3.2. Research Instruments

Four assessment instruments were employed in this research: the MTELP, a semi structured interview, a pre-test, and a post-test as explained below.

3.2.1. MTELP

The MTELP is a standardized test designed to assess the English language skills of non-native speakers. It is divided into three sections in a multiple-choice format. This dependable evaluation consists of 40 grammar items presented in a conversational style, 40 vocabulary questions on synonyms or sentence completion, and 20 items aimed at assessing reading comprehension. The entire test administration lasted for 100 minutes. While the MTELP is regarded highly reliable, its reliability was assessed using the KR-21 formula, which yielded a reliability index of .81 in the context of this study.

3.2.2. Semi-structured interview

The next instrument employed was a researcher-developed, semi-structured interview that focused on the participants' experiences with ChatGPT-based instruction and Microsoft Copilot. The research interview was researcher-made to fit the study's needs, using a semi-structured format. This approach balanced guided questions with the flexibility to explore emerging topics based on participants' responses, ensuring comprehensive and relevant data. The questions were specifically tailored to capture the unique perspectives of EFL learners using AI tools such as Microsoft Copilot and ChatGPT, enhancing the study's reliability and validity. The interview development process followed five stages. Ultimately, three main interview questions were crafted. To ensure the content validity of the interview, two experts in educational psychology were consulted to evaluate the questions. Based on their feedback, the researchers reduced the number of questions from seven to three.

1. We employ ChatGPT and Microsoft Copilot for the assessment of our writing skills and those of our colleagues. How would you differentiate between the two instruments and characterize your interaction with the instruments?
2. ChatGPT and Microsoft Copilot provide feedback on multiple elements of the writing, such as five subcategories of IMMs. How are the two instruments different regarding giving feedback?
3. What recommendations can you provide to enhance the effectiveness of utilizing ChatGPT and Microsoft Copilot in realizing IMMs?

3.2.3. Pretest

Prior to the intervention, a pretest was administered to assess the initial ability of participants to identify and use IMMs in their writing. The pretest included four samples of TOEFL iBT (Test of English as a Foreign Language) argumentative writing paragraphs. Participants were asked to identify and underline IMMs in these samples. This pretest helped to determine the participants' baseline proficiency in using IMMs. Participants who mistakenly underlined words with propositional meaning were randomly assigned to the study groups, while those who correctly identified IMMs were excluded from further analysis.

3.2.4. Posttest

At the end of the intervention, a posttest was conducted to evaluate the effectiveness of the instruction provided to each group. The posttest involved writing two argumentative paragraphs on topics taken from the TOEFL iBT. Participants were required to incorporate IMMs in their writing. The posttest assessed the improvement in participants' use of IMMs after the intervention. Two raters evaluated both the pretest and posttest, and the inter-rater reliability method yielded coefficients of 0.77 for the pretest and 0.81 for the posttest, which tend to be acceptable relaibilty indeces, as explained in Pallant (2020).

3.3. Intervention and Instructional Activities

The study incorporated a structured intervention over nine sessions spanning three weeks, with three sessions per week, each lasting two hours. Participants were divided into three groups: ChatGPT, Microsoft Copilot, and a control group. In the ChatGPT group, the initial sessions were dedicated to introducing participants to ChatGPT and providing thorough training on how to effectively use its functionalities. This included instructions on accessing the tool, navigating its interface, and understanding how to leverage its capabilities for writing assistance. Following these introductory sessions, participants in ChatGPT group engaged in extensive practice sessions where they wrote paragraphs and focused on improving their use of IMMs with the real-time feedback provided by ChatGPT. These practice sessions allowed participants to submit their paragraphs, receive instant feedback, and refine their use of IMMs, fostering a deeper engagement with the material.

Similarly, the Microsoft Copilot group began with sessions that familiarized participants with Microsoft Copilot's features and provided training on its use. Participants learned to draft, edit, and enhance their writing with real-time suggestions from Copilot, focusing on the strategic incorporation of IMMs. Throughout the practice sessions, participants wrote paragraphs, submitted them to Copilot, and received detailed recommendations for improving their writing.

Both experimental groups engaged with ten prompts per session, related to IMMs, and utilized their respective AI tools to receive guidance and feedback. In contrast, the control group adhered to conventional instruction methods, with participants practicing writing and receiving feedback from an instructor. These sessions focused on teaching the concepts of IMMs and their application in writing, relying on traditional methods where feedback was provided in subsequent sessions rather than in real-time.

Each group completed a total of nine sessions, with the final session dedicated to the posttest to assess their improved use of IMMs in argumentative writing. The structured approach in each group ensured comprehensive exposure to IMMs, with the experimental groups benefiting from the interactive and personalized feedback provided by the AI tools, while the control group followed a more traditional instructional method.

3.4. Procedure

The rationale for comparing ChatGPT-based instruction and Microsoft Copilot-based instruction is to understand the distinct capabilities and advantages each AI tool brings to language learning, particularly in helping EFL learners use IMMs in their argumentative writing. By evaluating the effectiveness of these tools, the study aims to determine which AI tool enhances learners' writing skills more effectively. Each chatbot uses different algorithms and feedback mechanisms, which could lead to varying impacts on the learners' writing. Including conventional instruction as a control allows us to measure the added benefits of AI-driven feedback compared to traditional methods. Additionally, examining learners' perceptions of ChatGPT and Microsoft Copilot provides insights into their usability and acceptance, which helps identify beneficial features and areas for improvement. Overall, this comparison aims to offer a comprehensive understanding of how various AI technologies can be integrated into language instruction to enhance educational.

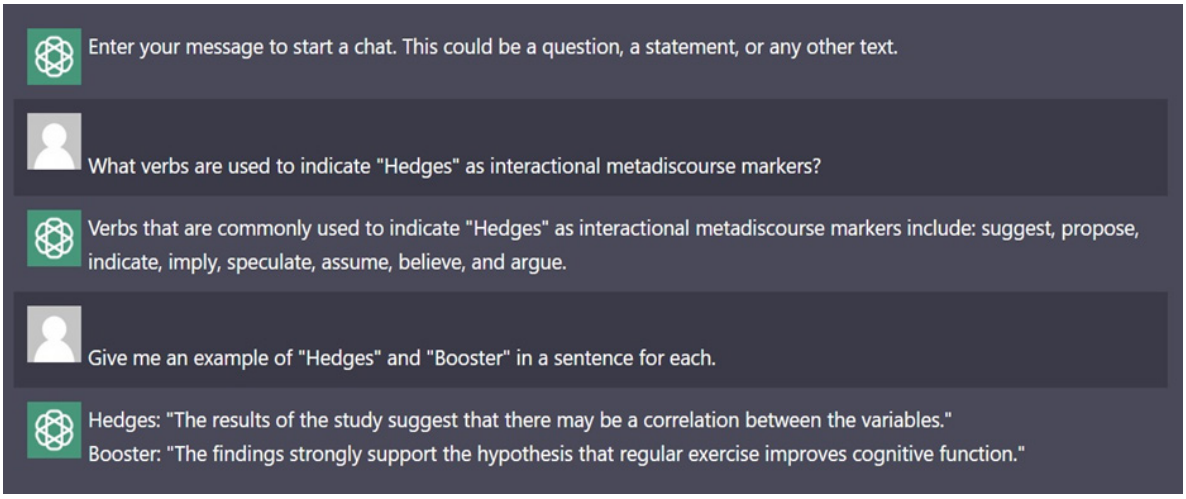
The present study followed the convergent parallel design (Mackey & Gass, 2023) through which the researchers simultaneously collected and examined both quantitative and qualitative data during the same phase of the study and later combined the two sets of findings to create a comprehensive interpretation. In order to evaluate the writing skills of the participants at the beginning and end of the study, the second researcher instructed students to compose an argumentative essay.

In the next phase, both experimental groups underwent a brief training program focused on utilizing AI-powered chatbots within the writing classroom. Participants in experimental groups (ChatGPT and Microsoft Copilot) and control group, each group consisting of thirty participants, were provided with printed files and papers focusing on introductory concepts of IMMs and subcategories of IMMs. Each experimental group benefitted from 15 personal computers (one for two participants) connected to the internet. Each session, the participants in both experimental groups received the same 10 prompts for IMMs realization to use ChatGPT and Microsoft Copilot, totaling for 60 prompts in six sessions. The last one session was for the posttest. Five sample prompts were as follows:

1. What verbs are used to indicate "Hedges" as IMMs?
2. Give me an example of "Hedges" and "Booster" in a sentence for each.
3. What are the most frequent forms of "Attitude markers" as an IMM?
4. Provide me with a paragraph including IMMs?
5. Determine the IMMs in the following paragraph.

Both ChatGPT and Microsoft Copilot groups used the same prompts and began to receive the answers from two AI-powered systems. As an example, Figure 1 shows the prompts and responses from ChatGPT. It should be noted that the teacher of the three groups was the second researcher.

FIGURE 1. ChatGPT response to prompts about IMMs



Accordingly, the same prompts were inserted into the Microsoft Copilot as shown in Figures 2 and 3.

FIGURE 2. Microsoft Copilot response to prompts about IMMs

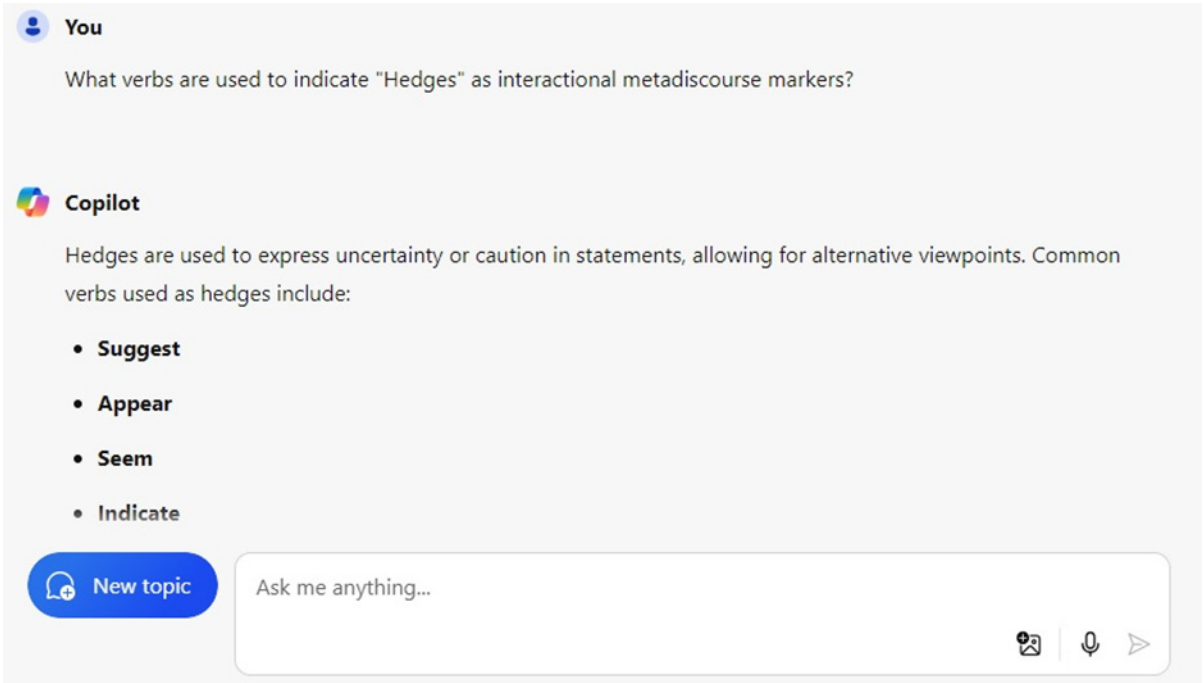
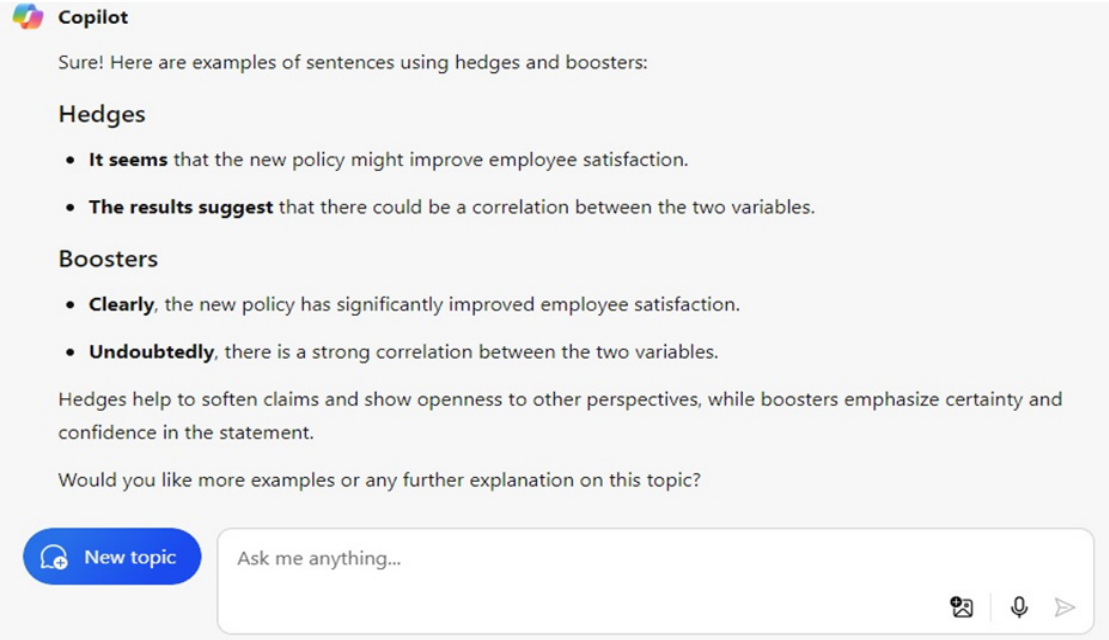


FIGURE 3. Microsoft Copilot response to prompts about IMMs



The control group received the conventional instruction, meaning that they were given some reading passages to identify IMMs with the help of the teacher. In the posttest, all participants in three groups were required to write two argumentative paragraphs (no more than 250 words each), on the following topics taken from TOEFL iBT.

1. Should dogs be allowed inside restaurants and stores?
2. Should schools give students money to invest in their futures?

Two raters evaluated both the pretest and post-test. The inter-rater reliability method between the two raters turned out to be 0.77 for the pretest and 0.81 for the post-test of the study. After treatment phase of the study, questions for interview were structured to gather insights from students regarding the use of ChatGPT and Microsoft Copilot. Participants in both experimental groups provided their viewpoints and opinions about ChatGPT and Microsoft Copilot.

3.5. Data Analysis and Coding Procedure

To answer the first research question, the researchers ran an analysis of covariance (ACNOVA) procedure to detect the variances between groups, due to the use of the pretest as the covatiate. Before running the statistical analysis, we checked the underlying assumptions and ensured they were satisfactorily met, as explained in the results section below.

As for the second research question, five participants were selected from each experimental (ChatGPT and Microsoft Copilot) group as volunteers to participate in the follow-up interview. Immediately after the posttest administration, those volunteers sat in a class and the second researcher began to ask them the interview questions, as listed in the research instruments section. The interview responses regarding learners'

perceptions of ChatGPT and Microsoft Copilot were thematically analyzed (Braun & Clarke, 2006). Inductive thematic analysis (Braun & Clarke, 2006) was chosen for analyzing interview data in this study as it provides a flexible, data-driven approach that allows themes to naturally emerge from participants' responses, ensuring an authentic and nuanced understanding of their experiences with AI tools. This method is particularly well-suited for exploring the complex perceptions and interactions of EFL learners, enhancing the credibility and depth of the findings, and aligning well with the exploratory nature of the research questions (Braun & Clarke, 2006). The interviews were audiotaped and then transcribed to be coded.

As shown in Table 2, a comprehensive guide was offered through the six phases of analysis for interview data, complemented by examples to illustrate the process (Braun & Clarke, 2006). It is essential to understand that qualitative analysis guidelines were not rigid rules but flexible principles that should be adapted to suit the specific research questions and data. Additionally, the analysis was not a straightforward progression from one phase to another but rather a recursive process, involving continuous back-and-forth movement through the phases as needed.

TABLE 2. Phases of inductive thematic analysis (Braun & Clarke, 2006, p. 87)

Phase	Description of the process	
1.	Familiarizing yourself with your data:	Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas.
2.	Generating initial codes:	Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3.	Searching for themes:	Collating codes into potential themes, gathering all data relevant to each potential theme.
4.	Reviewing themes:	Checking if the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic 'map' of the analysis.
5.	Defining and naming themes:	Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
6.	Producing the report:	The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.

The coding was done based on Adu's (2019) coding method. The researchers followed interpretation-focused coding in which they involved in a process of meaning-making, during which the generated codes reflected their comprehension and interpretation of the data. In employing this approach, the researchers drew upon their existing knowledge and experiences to influence the coding process. Given that the primary emphasis of the study was on IMMs, the researchers categorized codes under five subcategories of IMMs—hedges, boosters, engagement markers, attitude markers, and self-mentions. That is, when participants talked about their perceptions of utilization of ChatGPT and Microsoft Copilot, they focused on realizing hedges, boosters, attitude markers, self-mentions and engagement markers. Both researchers read the entire data and coded them indepenntly. When they disagreed on a code, a third coder, familiarized with qualitative data analysis and well versed in metadiscourse, was invited in to do the coding. The results of intercoder relaibilty, using Cronbach's alpha was .91. The rest of the discrepancies were discussed until full agrememnt was reached.

4. RESULTS

4.1. Investigation of the first research question

Regarding this research question, a one-way ANCOVA was performed to differentiate the results of the realization of IMMs in post-test for three groups, namely ChatGPT, Microsoft Copilot, and control group. Test of normality was conducted and data appeared to be normal for post-test (Pallant, 2020). Sig. value of Kolmogorov-Smirnova for posttest was checked (KS90 = .052, p > .05) (table 3).

To check for the assumption of equal variances, the data were evaluated. Levene's test result (F (2,87) = 1.217, p > .05) indicated that this assumption was upheld (Table 4).

Considering the assumption of linearity, the researchers found that a curvilinear relationship was not proven for each group. Therefore, linearity assumption was met (Figure 4).

TABLE 3. Tests of normality for post-test

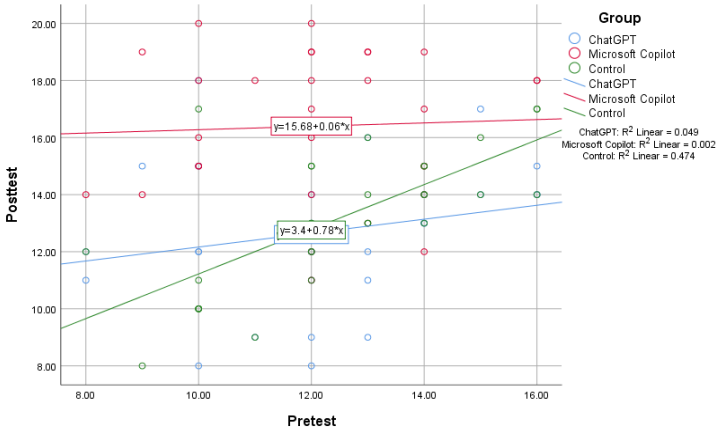
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Posttest	Statistic	df	Sig.	Statistic	df	Sig.
	.089	90	.076	.974	90	.065

TABLE 4. Levene's test of equality of error variances[a]

Dependent Variable: Posttest			
F	df1	df2	Sig.
1.217	2	87	.301

a. Design: Intercept + Pretest + Group

FIGURE 4. Linearity scatterplot for each group



The other assumption, homogeneity of regression slopes, was also satisfied (p > 0.5) (Table 5).

TABLE 5. Tests of between-subjects effects

Dependent Variable: Posttest						
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Corrected Model	339.577[a]	5	67.915	12.401	.000	.425
Intercept	236.964	1	236.964	43.270	.000	.340
Group	61.027	2	30.514	5.572	.005	.117
Pretest	51.837	1	51.837	9.465	.003	.101
Group * Pretest	36.473	2	18.237	3.330	.041	.073
Error	460.023	84	5.476			
Total	**18608.000**	**90**				
Corrected Total	**799.600**	**89**				

a. R Squared = .425 (Adjusted R Squared = .390)

TABLE 6. Descriptive statistics for IMMs in argumentative writing in the posttest

Dependent Variable: Posttest			
Group	**Mean**	**Std. Deviation**	**N**
ChatGPT	12.7000	2.52095	30
Microsoft Copilot	16.4000	2.58110	30
Control	13.1000	2.45441	30
Total	**14.0667**	**2.99738**	**90**

Table 6 displays an overview of the descriptive statistics for the scores across the three groups: ChatGPT group (M = 12.70, SD = 2.52), the Microsoft Copilot group (M = 16.40, SD = 2.58), the control group (M = 13.10, SD = 2.45). To decide the significance of these variations in the posttest, ANCOVA was employed.

The results of ANCOVA (Table 7) indicated that after adjusting for the initial differences on the pretest, statistically significant differences were seen among the groups on the posttest ($F(2, 86) = 22.23$, $p < .05$, partial eta squared = .34).

TABLE 7. ANCOVA post-test results on IMMs in argumentative writing across groups

Dependent Variable: Posttest						
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Corrected Model	303.103a	3	101.034	17.501	.000	.379
Intercept	234.935	1	234.935	40.694	.000	.321
Pretest	55.703	1	55.703	9.649	.003	.101
Group	256.698	2	128.349	22.232	.000	.341
Error	496.497	86	5.773			
Total	**18608.000**	**90**				
Corrected Total	**799.600**	**89**				

a. R Squared = .379 (Adjusted R Squared = .357)

To locate the differences between the three groups, pairwise comparisons were carried out (Table 8) to assess the significant differences in the realization of IMMs in argumentative writing across the three groups. The findings indicated that the difference was statistically significant ($p < .05$) in employing IMMs between the ChatGPT group and Microsoft Copilot group, with the latter demonstrating superior performance. Furthermore, the difference between ChatGPT group and control group was not statistically significant ($p > .05$). Overall, Microsoft Copilot group outperformed the other two groups in realizing IMMs in their argumentative writing.

TABLE 8. Pairwise comparisons for post-test scores of IMMs in argumentative writing across groups

Dependent Variable: Posttest						
Tukey HSD						
(I) Group	**(J) Group**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**	**95% Confidence Interval**	
					Lower Bound	**Upper Bound**
ChatGPT	Microsoft Copilot	-3.70000*	.65049	.000	-5.2511	-2.1489
	Control	-.40000	.65049	.812	-1.9511	1.1511
Microsoft Copilot	ChatGPT	3.70000*	.65049	.000	2.1489	5.2511
	Control	3.30000*	.65049	.000	1.7489	4.8511
Control	ChatGPT	.40000	.65049	.812	-1.1511	1.9511
	Microsoft Copilot	-3.30000*	.65049	.000	-4.8511	-1.7489

*. The mean difference is significant at the 0.05 level.

4.2. Investigation of the second research question

Grammarly and teacher feedback were two types of feedback that were given to the students in this study. On the one hand, Grammarly deals with three writing aspects: Diction, Language Use, and Mechanics. On the other hand, teacher feedback may cover the five writing aspects: Content, Organization, Diction, Language Use, and Mechanics. These types of feedbacks were given both in the first and second experimental groups.

The second research question sought to examine the learners' perception of the effectiveness of utilizing ChatGPT and Microsoft Copilot in realizing IMMs. The themes are presented in Table 9. The following two sample extracts show what the leaners in the experimental groups thought about using the chatbots. For example, in Extract 1, one of the language learners in the Microsoft Copilot group (participant 2) expresses his points of view concerning the several advantages Microsoft Copilot can offer and the ways it can be superior to ChatGPT.

Extract 1: The manner in which Microsoft Copilot assisted us in getting necessary information and organizing our thoughts about IMMs was quite remarkable. I had not previously recognized its potential as a metadiscourse learning companion. It's, in fact, superior to ChatGPT in a manner that it provided different choices, and you can also learn from them other than your own prompts. The best part of Microsoft Copilot was that the response was so clear and comprehensible. It delved into the details regarding the provided prompts.

In the second extract, another language learner in the ChatGPT group (participant 6) explained that although ChatGPT is helpful and supportive, it was not an accurate detector of the IMMs he was looking for.

Extract 2: I felt that it supported our writing and metadiscourse realization. However, the responses to the prompts were incomplete. It provided a short list of IMMs. More explanations were needed. Sometimes, misunderstanding happened. It could not find IMMs in a given paragraph very accurately. No more choices, other than the given prompts, were given through ChatGPT. I think the usual class by a teacher may be better for students to realize IMMs in argumentative writing.

TABLE 9. ChatGPT and Microsoft Copilot coded focus group data

Themes/sub-themes	AI-powered chatbot	Codes
Hedges	ChatGPT	**Positive perception** *Creation of a short definition* **Negative perception** *Less comprehensible and applicable* *No clear examples*
	Microsoft Copilot	**Positive perception** *Providing more ideas on the marker* *Clear definitions and responses* *Sensitive to implicit hedges* **Negative perception** *None*

Themes/sub-themes	AI-powered chatbot	Codes
Boosters	ChatGPT	**Positive perception** *Creation of strong examples* **Negative perception** *Less creativity in clarifying metadiscourse* *Examples in isolation not in a context*
	Microsoft Copilot	**Positive perception** *Introducing a reliable source* *Finding the marker in complex sentences* *Appropriate use of marker in a sentence* **Negative perception** *Promoting dependency on machine*
Attitude markers	ChatGPT	**Positive perception** *Good indicator of this marker* **Negative perception** *No ability to find the marker in a text* *Reduction of attention to the marker* *Lack of a complete concept of the marker*
	Microsoft Copilot	**Positive perception** *Promoting collaboration in realizing the marker* *Being focused on the marker* *Providing extra examples and prompts related to the marker* **Negative perception** *Giving lots of explanation and sometimes boring*
Self-mentions	ChatGPT	**Positive perception** *Creation of a clear definition* **Negative perception** *Unreliable source of marker detection* *Creation of confused examples in a sentence*
	Microsoft Copilot	**Positive perception** *Providing more explanation to make it compressible* *Being easy to understand the marker* *Providing argumentative paragraph with this marker* **Negative perception** *None*
Engagement markers	ChatGPT	**Positive perception** *Reliable source for this marker* **Negative perception** *Inappropriate use of the marker* *Reduction of attention to the marker* *Lack of a complete concept of the marker*
	Microsoft Copilot	**Positive perception** *Promoting motivation in realizing the marker* *Inspiring exercises* *Introducing more other prompts related to the marker* **Negative perception** *None*

Overall, the results of data taken from interview were in parallel with quantitative data analysis. Microsoft Copilot assisted the learners in employing and realizing IMMs in argumentative writing. Participants in Microsoft Copilot group showed their satisfaction. Participants highlighted the role of Microsoft Copilot in learning metadiscourse in general and IMMs in particular. They asserted that Microsoft Copilot gave them appropriate choices of IMMs in writing argumentative writing to make it convincing.

5. DISCUSSION

This section provides a comprehensive analysis and interpretation of the study's findings. First, we summarize the key results and their alignment with existing literature on the impact of AI-based tools on writing abilities. Then, we delve into a comparative analysis of ChatGPT and Microsoft Copilot, highlighting specific instances where one outperformed the other. We also explore the theoretical frameworks, as discussed in the review of lieterure, that support our findings, including feedback theory. Finally, we discuss the practical implications of our study for educators and learners, addressing the unique advantages and user experiences offered by Microsoft Copilot, and conclude with a consideration of the study's limitations and suggestions for future research.

While both ChatGPT and Microsoft Copilot exhibited enhanced employment of IMMs in writing abilities during the specified timeframe, the participants' performance in Microsoft Copilot group was markedly superior to that of their ChatGPT counterparts in the evaluation carried out after the treatment. The results align with those observed by scholars investigating the effect of AI-based tools on the writing abilities of university learners (Marzuki et al., 2023; Zhao, 2022). This study showed that Microsoft Copilot was better than ChatGPT in metadiscourse instruction, which is in conflict with Esfandiari and Allaf-Akbary's (2024b) study, resulting in the contribution of ChatGPT to the learners in realizing IMMs. Esfandiari and Allaf-Akbary found that ChatGPT-based instruction was more beneficial in employing IMMs in argumentative writing. However, it is widely accepted that AI-powered tools can strengthen language learning in terms of writing abilities (Kohnke et al., 2023; Rudolph et al., 2023). For example, learners in the Microsoft Copilot group reported that the feedback provided was detailed and helped them understand their mistakes better, and Microsoft Copilot made it easier to see where they needed to improve their writing. These ponits illustrate the positive perceptions of the tool's impact on the students' writing skills. Additionally, we have expanded the comparative analysis between ChatGPT and Microsoft Copilot. The detailed comparative analysis now includes specific instances where Microsoft Copilot outperformed ChatGPT, such as providing more contextual suggestions for improving IMMs.

Several factors underscoring the unique advantages and user experience offered by Microsoft Copilot are detected. One of the key aspects is Microsoft Copilot's seamless integration with Microsoft Office applications like Word and Outlook, providing a familiar and user-friendly interface that students are already accustomed to, which reduces the learning curve and makes the tool more accessible (Scholl, 2024). This familiarity contributes to higher student preference and engagement. Additionally, Microsoft Copilot's advanced algorithms deliver contextually relevant suggestions and feedback, which help students effectively use IM by suggesting appropriate transitions, hedges, and boosters that enhance the persuasiveness and clarity of their arguments (Esfandiari & Allaf-Akbary, 2024c). This targeted feedback aligns closely with the principles of effective feedback in language learning, as outlined by feedback theory, which emphasizes the importance of specific, timely, constructive, and actionable feedback in improving performance. Another significant advantage of Microsoft Copilot is its ability to provide real-time writing assistance, allowing students to receive immediate suggestions and corrections as they draft their essays. This real-time support is crucial in academic settings where time is often limited, and immediate feedback can significantly enhance the writing process (Campesato, 2024). Moreover, Microsoft Copilot's customization features enable

students to tailor the tool to their specific needs, providing a more personalized and effective writing assistant compared to more generic tools like ChatGPT. The positive user experience with Microsoft Copilot, combined with its ability to provide feedback that meets educational standards and expectations, further solidifies its preference among students. Additionally, Microsoft's reputation for strong security and data privacy may instill more confidence in students and educational institutions, contributing to its widespread adoption (Miller, 2024).

The outcomes derived from the investigation of interview data revealed a generally favorable perception considering the influence of AI-based tools. This observation aligns with existing literature concerning learners' viewpoints on the effects of AI-based tools on their writing practices (Yan, 2023). The generally favorable disposition can be attributed to the excitement surrounding the execution of AI chatbots within academic contexts. The results regarding the role of Microsoft Copilot in facilitating content generation are consistent with the claims made by Guo et al. (2022) and Marzuki et al. (2023). The results of our study indicated that a significant number of learners held positive views regarding Microsoft Copilot. This supports the results of Chen et al. (2020), which spotlighted educators' recognition of the difficulties in learning and their openness to integrating AI-enhanced learning methodologies in their future instructional practices. Microsoft Copilot presents significant advantages for individuals engaged in L2 learning by facilitating the generation of precise and contextually relevant points. Through its provision of instantaneous suggestions and corrections, Microsoft Copilot aid learners in refining their linguistic abilities while deepening their comprehension of grammar and lexicon in a foreign language (Pym & Hao, 2024). Furthermore, the tools can deliver customized recommendations aligned with user preferences, thereby optimizing the learning experience and catering to specific requirements.

As the findings of the study show, Microsoft Copilot stands as a crucial resource for L2 learners aspiring to elevate their language proficiency. Microsoft Copilot serves as a valuable tool for helping EFL leaners learn IMMs in an L2 by delivering immediate suggestions and corrections during the writing process (Pan, 2024). This functionality enables users to broaden their knowledge of IMMs by presenting alternative terms or expressions in the desired language (Khadimally, 2022). Furthermore, Microsoft Copilot can offer metadiscourse recommendations that are contextually pertinent to the material being composed, thereby facilitating the learning of new IMMs within the correct context (Bowen & Watson, 2024). By leveraging the language support capabilities of Copilot, users can refine their metadiscourse realization skills and elevate their overall competence in an L2. Through consideration of prior knowledge and experience, AI can select tasks, activities and plan lesson in language learning (Vurdien & Chambers, 2024).

The fidnings can be explained in terms of feedback theory, which emphasizes the importance of providing specific, timely, and actionable feedback to improve performance. In our study, the AI chatbots, particularly Microsoft Copilot, provided immediate and detailed feedback on the use of IMMs. This immediate feedback acted as a critical support mechanism, helping learners internalize and apply IMMs effectively in their writing. The superior performance of the Microsoft Copilot group can be attributed to its ability to deliver precise and contextually relevant feedback. This aligns with feedback theory, which posits that effective feedback should be clear, specific, and given promptly to facilitate learning and improvement (Nassaji & Kartchava, 2024). The positive perceptions reported in semi-structured interviews suggest that learners found the detailed feed-

back from Microsoft Copilot highly beneficial, reinforcing their understanding and application of IMMs. The structured feedback helped learners focus on key aspects of their writing, making the necessary adjustments in real-time, thus enhancing their overall writing skills and proficiency in using IMMs.

The observed differences in the effectiveness of IMMs between Microsoft Copilot and Chat-GPT may be attributed to their underlying transformer model architectures. Microsoft Copilot is built on a sophisticated transformer model specifically designed to enhance language understanding and generation, optimized for tasks requiring nuanced language use, such as incorporating IMMs. This model effectively handles long-range dependencies and contextualizes information across sentences, generating coherent and contextually relevant IMMs (Esfandiari & Allaf-Akbary, 2024c). In contrast, ChatGPT, based on the advanced GPT-4 architecture, is a generalist model excelling in generating coherent texts across diverse topics but may not prioritize specific linguistic features like IMMs as effectively as specialized models (Roumeliotis & Tselikas, 2023). The composition of training data and frequency of updates also significantly impact the language generation capabilities of these AI models. Microsoft Copilot benefits from training on datasets rich in academic and argumentative texts, enhancing its ability to generate texts with effective IMMs. In contrast, Chat-GPT's broader corpus may not prioritize these features as strongly. Additionally, Microsoft Copilot may receive more frequent updates focused on the latest academic and professional writing trends, further refining its capabilities (Minnick, 2025). Algorithmic enhancements, such as advanced techniques for improved contextual understanding and discourse modeling, contribute to Microsoft Copilot's superior performance. Sophisticated feedback mechanisms within Copilot allow it to learn from user interactions, continuously refining its language generation capabilities over time (Stratton, 2024). These combined factors—model architecture, specialized training data, frequent updates, and algorithmic enhancements—may explain the superior performance of Microsoft Copilot in generating effective and contextually relevant IMMs compared to Chat-GPT.

Overall, the synergy between feedback theory and the practical application of GPT as a writing assistant, as exemplified by Microsoft Copilot, highlights the transformative potential of integrating effective feedback mechanisms with advanced AI tools to enhance language learning and writing processes, ultimately supporting students in achieving their academic goals.

5.1. Implications, Limitations, and Suggestions for Further Research

The results of this study have several implications for future teaching practices, particularly in the context of teaching metadiscourse in argumentative writing. Language educators can integrate AI tools like Microsoft Copilot into their curriculum to provide personalized feedback on the use of IMMs, helping learners understand the importance of these markers in structuring their arguments and engaging their audience. The use of AI in language instruction can also help identify areas where students struggle with metadiscourse and offer targeted interventions to address these challenges. Additionally, the educators should consider providing training to students on how to effectively use AI tools, ensuring that they can fully leverage the benefits of these technologies in their learning process. By incorporating AI-powered feedback, educators can enhance students' awareness and usage of IMMs, ultimately improving the overall quality of their argumentative writing.

This research posits that AI-driven language learning can serve a vital function in fostering the educational experiences of EFL learners, thereby significantly advancing their language learning processes. These tools provide a more delightful and customized learning environment, which fosters and enthusiasm for ongoing language study. Platforms enhanced by AI, such as Microsoft Copilot, enable learners to apply language abilities at their own convenience and from various locations. This accessibility empowers learners to assume control over their educational trajectories, promoting a feeling of independence and self-control. Furthermore, the study emphasizes the great influence of AI-based language learning within EFL classrooms, particularly their capacity to create personalized and adaptive learning experiences. Interacting with these dynamic AI systems allows learners to receive immediate feedback, constructive assessments, and diverse sentence structures, all of which contribute to continuous improvement and enhance their confidence in using an L2 language.

By offering personalized feedback, these AI tools enable tailored instructional strategies, addressing specific learner needs and fostering critical writing skills through real-time feedback. Additionally, they streamline the assessment process, reducing educators' workloads and allowing focus on interactive teaching. The integration of AI chatbots creates adaptive learning environments that offer continuous support, improving student motivation and outcomes. Furthermore, AI-driven insights aid educators' professional development by identifying areas needing improvement, leading to more effective teaching practices. Overall, AI chatbots can revolutionize language instruction through personalized learning, efficient assessment, and continuous professional development.

This study is, however, susceptible to several limitations, including the use of a restricted sample size, the absence of a delayed posttest to assess learners' understanding of IMMs, and a focus solely on IMMs rather than encompassing interactive metadiscourse as well. Additionally, all participants were drawn from a pool of individuals possessing advanced language proficiency. The research employed an interview to collect data regarding participants' attitudes towards ChatGPT and Microsoft Copilot. Other methodologies, such as think-aloud protocols, could be employed to obtain more profound insights into learners' perspectives.

6. CONCLUSIONS

The findings from the present study suggested that Microsoft Copilot may significantly contribute to improving IMMs in writing argumentative paragraphs. The results indicated that Microsoft Copilot group demonstrated higher levels of comprehension of IMMs in the posttest. The notable effects, along with learners' positive views regarding these tools, can contribute a new dimension to the existing literature on AI-powered chatbots. The findings of the present study enhance and expand upon theories regarding feedback as a dialogic procedure, positioning Microsoft Copilot and ChatGPT as feedback applications that may be seamlessly incorporated into extensive writing courses. The results of the research indicated that the communication enabled by the chatbots throughout the writing process significantly enhance learners' learning. This engagement positively affects learners' approaches to seeking feedback, their responsiveness to it, and their ability to enhance their argumentative writing skills.

Based on the findings of the study, learner autonomy is one área AI tools such as copilot paly a key role in developing it. EFL learners can use them at their own pace to learn new language points, ask them to provide individulaised feedback, and adjust their learning following the feedback they receive. These processes will help learners to take resposibilty for their own learning, become accountable, and eventually achieve autonomy in learning. We recommend language teachers encourage language leanrers to use such tools and familairze them with the most frequently used AI tools, because teachers will not always be available to help them. Another área the results of the study may be useful is intelligent data-driven learning (Esfandiari & Allaf-Akbary, 2024c), because AI tools such as copilot are inteligent and based on large language moldes (LLMs) and can be particularly helpful in fostering writing abilities. EFL learners can use these tools to provide them with several authentic senstence examples, discover the rules by analysing such senetnces, and learn how a particular rule is naturally used in real-life sitautions. Language teachers can also utilize such tools to provide learners with sufficient natural input, exposing them to the language native-English speakers actually use.

Declaration of conflicting interest

There is no conflict of interest in this work.

7. FUNDING

This research received no external funding.

8. REFERENCES

Abdelrahim A. A. M., & Abdelrahim, M. A. M. (2020). Teaching and assessing metadiscoursal features in argumentative writing: A professional development training for EFL teachers. *International Journal of Applied Linguistics, 30*, 70-91. https://doi.org/10.1111/ijal.12264

Ädel, A. (2023). Adopting a 'move' rather than a 'marker' approach to metadiscourse: A taxonomy for spoken student presentations, *English for Specific Purposes, 6*(69), 4-18. https://doi.org/10.1016/j.esp.2022.09.001

Adu, P. (2019). *A step-by-step guide to qualitative data coding.* Taylor & Francis.

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix, 2*(1), 41–49. https://doi.org/10.56540/jesaf.v2

Allagui, B. (2023). Chatbot Feedback on students' writing: Typology of comments and effectiveness. In O. Gervasi, B. Murgante, A. M. A. C. Rocha, C. Garau, F. Scorza, Y. Karaca, & C. M. Torre (Eds.), *International conference on computational science and its applications* (pp. 377–384). Springer. https://doi.org/10.1007/978-3-031-37129-5_31

Amirjalili, F., Neysani, M., & Nikbakht. A. (2024). Exploring the boundaries of authorship: a comparative analysis of AI-generated text and human academic writing in English literature. *Frontiers in Education, 9*(2), 76-98. https://doi.org/10.3389/feduc.2024.1347421

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, 100745. https://doi.org/10.1016/j.asw.2023.100745

Bautista, J. (2024). *Chat GPT for students: How to become an A+ student using Chat GPT.* Benjamin Bautista.

Bowen, J. A., & Watson, C. E. (2024). *Teaching with AI: a practical guide to a new era of human learning*. Johns Hopkins University Press.

Braun, V., & Clarke, V. (2006) Using thematic analysis in psychology, *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Campesato, O. (2024). *GPT-4 for developers*. Mercury Learning and Information.

Chen, X., Xie, H., Zou, D., & Hwang, G. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*(4), 100002. https://doi.org/10.1016/j.caeai.2020.100002

Clark, V. E. (2024). *First language acquisition*. Cambridge University Press.

Dai, W., Lin, J., Jin, F., Li, T., Tsai, Y., Gasevic, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. *IEEE International Conference on Advanced Learning Technologies*. https://doi.org/10.35542/osf.io/hcgzj

El-Dakhs, D. A. S., Yahya, N., & Pawlak, M. (2022). Investigating the impact of explicit and implicit instruction on the use of interactional metadiscourse markers. *Asian-Pacific Journal of Second and Foreign Language Education, 44*(7), 2-21. https://doi.org/10.1186/s40862-022-00175-0

Escalante, J., Pack, A. & Barrett, A. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20*, Article 57. https://doi.org/10.1186/s41239-023-00425-2

Esfandiari, R., & Allaf-Akbary, O. (2024a). The role of learning-oriented language assessment in promoting interactional metadiscourse in ectenic and synoptic EFL learners. *Journal of Modern Research in English Language Studies, 11*(3), 181-206. https://doi:10.30479/jmrels.2024.19777.2305

Esfandiari, R., & Allaf-Akbary, O. (2024b). The role of ChatGPT-based instruction and flipped language learning in metadiscourse use in EFL learners' argumentative writing and their perceptions of the two instructional methods. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 43*(3), 45-65. https://doi: 10.22099/tesl.2024.49975.3277

Esfandiari, R., & Allaf-Akbary, O. (2024c). Assessing interactional metadiscourse in EFL writing through intelligent data-driven learning: the Microsoft Copilot in the spotlight. *Language Testing in Asia, 14*, Article 51. https://doi.org/10.1186/s40468-024-00326-9

Esfandiari, R., & Khatibi, Z. (2022). A cross-cultural, cross-contextual study of interactional metadiscourse in academic research articles: An interpersonal approach. *Southern African Linguistics and Applied Language Studies, 40*(2), 227-242, https://doi: 10.2989/16073614.2022.2056064

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *NPJ Digital Medicine, 75*(6), 41-57. https://doi.org/10.1038/s41746-023-00819-6

Gordon, C. (2024). *Intertextuality 2.0: Metadiscourse and meaning-making in an online community*. Oxford University Press.

Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing, 54*, 100666. https://doi.org/10.1016/j.asw.2022.100666

Ho, V., & Li, C. (2018). The use of metadiscourse and persuasion: An analysis of first year university students' timed argumentative essays. *Journal of English for Academic Purposes, 33*(2), 53-68. https://doi.org/10.1016/j.jeap.2018.02.001

Holmes, W., Persson, J., Chounta, I. A., Wasson, B., & Dimitrova, V. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law*. Council of Europe.

Huang, A. Y. Q., Lu, O. H. T., & Yang, S. J. H. (2023). Effects of artificial intelligence–enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers & Education, 194*(2), 104684. https://doi.org/10.1016/j.compedu.2022.104684

Huang, F., & Zou, B. (2024). English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI, and innovativeness. *Computers in Human Behavior, 159*(2), 108355. https://doi.org/10.1016/j.chb.2024.108355

Huang, Q., Liu, Y., Chen, S., & Zhao, M. (2024). Generative AI acceptance, teachers' enthusiasm, and learners' self-efficacy: Impact on EFL learners' well-being. *Journal of Language*

and Education Research, 38, 75-89. https://doi.org/10.1111/ejed.12770

Hyland, K. (2019). Metadiscourse*: Exploring interaction in writing* (2nd edition). Continuum.

Hyland, K., & Jiang, F. K. (2022). Metadiscourse choices in EAP: An intra-journal study of JEAP. *Journal of English for Academic Purposes, 60*(3), 101-165. https://doi.org/10.1016/j.jeap.2022.101165

Khadimally, S. (2022). *Applications of machine learning and artificial intelligence in education*. IGI Global.

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies, 27*(5), 6069–6104. https://doi.org/10.1007/s10639-021-10831-6

Kim, J., Yu, S., Detrick, R., & Li, N. (2024). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies, 2*(34), 1-36 https://doi.org/10.1007/s10639-024-12878-7

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537-550. https://doi.org/10.1177/00336882231162868

LaRue Keeley, K. (2024). *AI Applications and Strategies in Teacher Education*. Information Science Reference. https://doi.org/10.4018/979-8-3693-5443-8

Lee, S. S., Li, N., & Kim, J. (2023). Conceptual model for Mexican teachers' adoption of learning analytics systems: The integration of the information system success model and the technology acceptance model. *Education and Information Technologies*, 1–26. https://doi.org/10.1007/ s10639-023-12371-7

Li, Z., & Xu, J. (2020). Reflexive metadiscourse in Chinese and English sociology research article introductions and discussions. *Journal of Pragmatics, 159*(2), 47-59. https://doi.org/10.1016/j.pragma.2020.02.003

Lim, G. S. (2011). The development and maintenance of rating quality in performance writing assessment: A longitudinal study of new and experienced raters. *Language Testing, 28*(4), 543-560.

Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & 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, 11*(50), 1-24. https://doi.org/10.1186/s40561-024-00342-5

Mackey, A. (2020). *Interaction, feedback and task research in second language learning methods and design*. Cambridge University Press.

Mackey, A., & Gass, S. M. (2023). *Current approaches in second language acquisition research: A practical guide*. Wiley.

Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environment, 11*, 9. https://doi.org/10.1186/s40561-024-00295-9

Mahapatra, S. K. (2021). Online formative assessment and feedback practices of ESL teachers in India, Bangladesh and Nepal: A multiple case study. *Asia-Pacific Education Researcher, 30*(6), 519–530. https://doi.org/10.1007/s40299-021-00603-8

Marzuki, S., Widiati, U., Rusdin, D., Darwin, R., & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education, 10*(2), 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Mathur, M., & Mahapatra, S. (2022). Impact of ePortfolio assessment as an instructional strategy on students' academic speaking skills: An experimental study. *Computer-Assisted Language Learning Electronic Journal, 23*(3), 1–23.

McCourt, D. M. (2023). *The new constructivism in international relations theory*. Bristol University Press.

Miller, M. (2024). *Using artificial intelligence: Absolute beginners guide*. Pearson Education INC.

Minnick, C. (2025). *Microsoft Copilot for Dummies*. John Wiley & Sons, Inc.

Moser, A. (2020). *Written corrective feedback: The role of learner engagement: A practical approach*. Springer International Publishing.

Nassaji, H., & Kartchava, E. (2024). *The Cambridge handbook of corrective feedback in second language learning and teaching*. Cambridge University Press.

Pallant, J. (2020). *SPSS Survival Manual: A step by step guide to data analysis using IBM SPSS* (7th ed.). Routledge.

Pan, F. (2024). *AI in language teaching, learning, and assessment*. IGI Global.

Pearson, S. W., & Abdollahzadeh, E. (2023). Metadiscourse in academic writing: A systematic review. *Lingua, 293*(9), 103-561. https://doi.org/10.1016/j.lingua.2023.103561

Phakiti, A. (2003). A closer look at gender and strategy use in L2 reading. *Language Learning, 53*(4), 649-702. https://doi.org/10.1046/j.1467-9922.2003.00239.x

Pym, A., & Hao, Y. (2024). *How to augment language skills: Generative ai and machine translation in language learning and translator training*. Taylor & Francis.

Qiu, X., Wang, Y., Dartey, E. A., & Kim, M. (2024). Interactional metadiscourse in expert and student disciplinary writing: Exploring intrageneric and functional variation. *English for Specific Purposes, 73*(5), 124-140. https://doi.org/10.1016/j.esp.2023.10.007

Rostami Aboo saeedi, A. A., Hasanpour Ghaleh, M., Kheiri, S., & Mohammadi, S. (2023). A qualitative account of writers' perceptions of thesis writing in applied linguistics: The case of interactional meta-discourse markers. *Iranian Journal of English for Academic Purposes, 12*(2), 31-48.

Roumeliotis, K. I., & Tselikas, D. N. (2023). ChatGPT and Open-AI Models: A Preliminary Review, *Future Internet, 15*(6), 192. https://doi.org/10.3390/fi15060192

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

Sadiku., M. N. O., Musa, M. S., & Chukwu, U. C. (2022). *Artificial Intelligence in Education*. iUniverse.

Scholl, H. (2024). *Microsoft CoPilot AI expertise*. Publisher s21598.

Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.

Stratton, J. (2024). *Copilot for Microsoft 365: Harness the power of generative AI in the Microsoft apps you use every day*. Apress.

Vurdien, R., & Chambers, W. (2024). *Technology-mediated language learning and teaching*. IGI Global.

Wang, M., & Guo, W. (2023). *The potential impact of ChatGPT on education: Using history as a rearview mirror*. ECNU Review Education. https://doi.org/10.1177/20965311231189826

Winstone, N. E., & Carless, D. (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. Routledge.

Wu, X., Zhang, Y., Li, J., & Wang, L. (2024). Exploring the factors influencing EFL college learners' willingness to use AI tools in distributed learning environments. *Journal of Educational Technology & Society, 13*, 45-59. https://doi.org/10.3389/fpsyg.2022.870777

Xue, Y. (2024). Towards automated writing evaluation: A comprehensive review with bibliometric, scientometric, and meta-analytic approaches. *Education and Information Technologies, 29*, 19553–19594. https://doi.org/10.1007/s10639-024-12596-0

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Educational Information Technology, 28*(2), 13943–13967 https://doi.org/10.1007/s10639-023-11742-4

Zhao, X. (2022). Leveraging artificial intelligence (AI) technology for English writing: Introducing word tune as a digital writing assistant for EFL writers. *RELC Journal, 54*(3), 65-89. https://doi.org/10.1177/00336882221094089

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 74-91 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20529

Un chatbot apoyando la enseñanza de programación de computadoras

A chatbot supporting computer programming

RECIBIDO 19/09/2024 ACEPTADO 03/12/2024 PUBLICADO 01/06/2025

Omar Ayala Cadena
ORCID 0009-0009-2942-2470
Universidad Autónoma del Estado de México, México
oayalac426@alumno.uaemex.mx

Irene Aguilar Juárez
ORCID 0000-0003-4747-0336
Universidad Autónoma del Estado de México, México
iaguilarj@uaemex.mx

Joel Ayala de la Vega
ORCID 0000-0003-3279-4143
Universidad Autónoma del Estado de México, México
jayalad@uaemex.mx

Adriana Bustamante Almaraz
ORCID 0000-0002-5476-4140
Universidad Autónoma del Estado de México, México
abustamantea@uaemex.mx

RESUMEN

En el ámbito educativo se requieren de herramientas tecnológicas que tengan disponibilidad inmediata en cualquier momento del día, por lo que los asistentes conversacionales son la herramienta ideal para apoyar el proceso de aprendizaje del estudiantado facilitando la resolución de dudas y acceso a información especializada. Este trabajo describe el desarrollo de un chatbot en Dialogflow plataforma de Google integrado con la aplicación de mensajería Telegram para asistir el aprendizaje de programación de computadoras. Se abordan los chatbots desde su definición, antecedentes y aplicaciones en la educación con el objetivo de desarrollar un chatbot específico para la comunidad del Centro universitario UAEM Texcoco y ser usado como un medio de consulta de información y atención a dudas. Se utilizó el modelo de desarrollo espiral, considerando el léxico utilizado por el estudiantado mismo que fue analizado con Lancsbox, se realizaron tres instrumentos de recolección de datos y se validaron con pruebas estadísticas, la primera prueba es el Coeficiente de Alfa de Cronbach esta se aplicó al cuestionario aplicado tanto a docentes como a estudiantes para evaluar la opinión de usabilidad y funcionamiento del agente conversacional. Se plantearon cuatro hipótesis de investigación, que fueron evaluadas con la segunda prueba estadística; la T de Student. Se comprobaron las hipótesis con los datos de las muestras y los resultados obtenidos permiten aceptar las hipótesis formuladas por lo que se concluye que el asistente conversacional es percibido como una herramienta que permite apoyar el aprendizaje de la programación de computadoras. Los chatbots son un ejemplo de aplicación de nuevas herramientas en el ámbito educativo, el uso de este asistente permite la interacción del

estudiante como si estuvieran en asesorías con los profesores, de modo que el asistente es una herramienta innovadora personalizada como medio de consulta de información dirigida a un aprendizaje de temas especializados.

PALABRAS CLAVE Inteligencia Artificial; Educación; Tecnología Educativa; Universidad; Motivación para el aprendizaje.

ABSTRACT

In the educational sector, technological tools that are immediately available at any time of the day are required, so conversational assistants are the ideal tool to support the learning process of students by helping them to resolve doubts and access specialized information. This project describes the creation of a chatbot in Google's Dialogflow platform integrated with the Telegram messaging application to assist in the learning of computer programming. Chatbots are approached from their definition, history, and applications in education with the purpose of developing a specific chatbot for the community of the UAEM Texcoco University Center to be used as a tool for consulting information and answering questions. The spiral model of design was used, considering the vocabulary used by the students, which was analyzed with Lancsbox, three instrument for data collection were carried out and validated with statistical tests, the first test is the Cronbach's Alpha Coefficient, which was applied to the test applied to both teachers and students to evaluate the opinion of usability and performance of the conversational agent. Four research hypotheses were raised, which were evaluated with the second statistical test: Student's - T test. The hypotheses were tested with the sample data and the results obtained allow accepting the hypotheses formulated, so it is concluded that the conversational assistant is perceived as a tool that provides support for learning computer programming. The chatbots are an example of application of new tools in the educational sector, the use of this assistant makes possible the interaction of the student as if they were in tutorials with teachers, so that the assistant is an innovative personalized application as a reference of information oriented to the learning of specialized topics.

KEYWORDS Artificial Intelligence; Education; Educational Technology; University; Learning Motivation.

1. INTRODUCCIÓN

Recientemente en el sector educativo se están explorando e implementando nuevas herramientas tecnológicas con la finalidad de mejorar los resultados de la acción educativa. De las más recientes los chatbots o asistentes conversacionales, representan una nueva forma de dirigir y ayudar en el aprendizaje, facilitando la personalización del formato de preguntas frecuentes mediante respuestas cortas y concretas que refuerzan el aprendizaje de las y los estudiantes.

Los asistentes conversacionales son aplicaciones tecnológicas que facilitan la entrega de la información de manera inmediata, para ser una vía que fomenta el autoaprendizaje (Arredondo Castillo, 2021), además los chatbots sirven de apoyo a los docentes dado que ayuda a delegar algunas cargas de trabajo como es la atención al estudiantado para solventar dudas, permitiéndoles dedicar su tiempo a otras actividades que tienen dentro de sus planeaciones escolares.

Los asistentes conversacionales de acuerdo con Villón Cabrera (2020) son programas que se comunican en lenguaje natural mediante texto imitando una conversación humana, Bagchi (2020) por su parte los considera como un programa informático que imita y produce conversaciones inteligentes con el uso de la computación capaces de simular y mantener una conversación con las personas.

La característica más importante de estos asistentes es estar dirigidas a realizar ciertas tareas, siguiendo un flujo de dialogo basado en una estructura predefinida y programada; Astruga (2021) los considera como una máquina de estados finita.

Sin embargo, aunque hay una difusión masiva de chatbots por el avance de estas tecnologías en el mercado, sus comienzos datan de los años sesenta siendo uno de los primeros chatbots ELIZA que fue capaz de simular una conversación entre un psicoanalista y un paciente (Thorat y Jadhav, 2020), seguido de Parry que fue un intento de modelar el comportamiento de la esquizofrenia paranoide como describe Deshpande y Chandak (2022). ALICE por su parte contaba con la habilidad de obtener información de cualquier tema en la web y crear una conversación de cualquier tema con los usuarios (Adamopoulou y Moussiades, 2020). Llegando hasta el día de hoy, donde existen los asistentes personales de voz de grandes compañías como son Siri de Apple, Cortana de Microsoft o Alexa de Amazon como mencionan Adamopoulou y Moussiades (2020).

La clasificación que se hace de los asistentes conversacionales de acuerdo con Vishwakarma y Pandey (2021) son aquellos que se basan en objetivos, en conocimiento, en servicio y aquellos que generan una respuesta. Por su lado Deshpande y Chandak (2022) los clasifican como chatbots basados en reglas, basados en Inteligencia Artificial e híbridos. Finalmente, Thorat y Jadhav (2020) identifica igualmente los basados en reglas y añaden los de autoaprendizaje.

Los chatbots han sido considerados en el ámbito educativo con algunas de sus aplicaciones:

a. Como tutores que ayudan al proceso de aprendizaje, que pueden adaptarse para seleccionar y secuenciar contenidos que el estudiantado requiere para continuar el avance de su propio aprendizaje, los chatbots ayudan a personalizar el aprendizaje, mejorando la metacognición y aumentando su motivación.
b. Como programas de práctica donde el objetivo es incrementar las habilidades del estudiantado, los cuales funcionan por medio de preguntas al alumnado o presentación de problemas de tal forma que el chatbot evalúa las respuestas para entregar una retroalimentación inmediata (García Brustenga et al., 2018).

Considerando que los chatbots permiten ser de utilidad en el ámbito educativo por sus aplicaciones existen diversas plataformas que apoyan la implementación de estas herramientas tecnológicas; Deshpande y Chandak, (2022), García Brustenga et al. (2018), , Maher et al. (2020), Varela Tapia et al. (2022) y Vishwakarma y Pandey (2021) consideran las siguientes plataformas existentes en el mercado para crearlos entre ellos: Watson de IBM, Dialogflow de Google, LUIS de Microsoft, Lex de Amazon, Wit.AI de Facebook respaldados por grandes industrias y otros como Chatfuel, Pandorabots, ChatScript o ChatterBot que son producto de empresas con menos recursos.

Para generar una experiencia de uso de este tipo de herramientas y lograr un aprendizaje centralizado es necesario vincularlos con aplicaciones de mensajería externas a las plataformas en donde se desarrollan los chatbots, como pueden ser Messenger de Facebook, Telegram, Slack, Skype, Twitter, entre otros más.

1.1. Justificación

La creciente evolución del software ha permitido que los escenarios educativos sean muy dinámicos y retadores; en el estudio de Cervantes y Chaparro (2021) se evidencia que, derivado de la pandemia COVID-19, se presentó un incremento en el uso del internet, en el uso de algunas aplicaciones y en las actividades en línea, en conclusión, se generaron nuevos procesos de interacción social y el uso de las nuevas aplicaciones se han convertido en parte de la cotidianidad. En México la pandemia del año 2020 aceleró el uso del software educativo en todos los niveles escolares. Los docentes se vieron obligados a usar software desconocido para ellos y surgió la necesidad de seleccionar el software, evaluar los resultados obtenidos y acompañar virtualmente el aprendizaje de los estudiantes en condiciones muy limitadas (García Martínez y Silva Payró, 2022; Ramos Mendoza et al., 2023). Las actividades presenciales tuvieron que ser reemplazadas por el uso de las plataformas educativas y las asesorías fueron difíciles de ofrecer bajo las normativas estrictas de confinamiento (Berumen López et al., 2023; Carrazco Zendejas, 2021; Cujia Berrío, 2023). Al mismo tiempo los docentes invirtieron sus esfuerzos en aprender a usar software educativo de forma obligada y el tiempo de atención a las dudas del estudiantado se redujo al mínimo (Ortega Sánchez, 2021).

El cambio que más afectó a la población estudiantil se observó en la interacción alumno – profesor; pues se minimizó el contacto presencial y la disponibilidad de tener asesorías escolares (Mendoza Navarro, 2022; Morales Bonilla y Bustamante Peralta, 2021; Ponce Ceballos et al., 2022). Si bien hoy en día esta contingencia ya no es vigente, la realidad es que se transformó de forma definitiva el escenario de las asesorías, pues la Universidad Autónoma del Estado de México se adaptó a generalizar en los programas de estudio la modalidad virtual o mixta, por lo que se generaron nuevas necesidades en los actores educativos ya que ambas partes se enfrentaron al reto de interactuar en múltiples modalidades con poca disponibilidad de espacios para reuniones presenciales.

En este contexto es importante analizar desde el enfoque didáctico los resultados de la implementación de los asistentes virtuales como apoyo en las asesorías escolares; por otro lado, los asistentes conversacionales promueven el aprendizaje activo en las y los jóvenes y por eso es necesario explorar los beneficios de este tipo de software.

El asistente conversacional genera una interacción personalizada de acuerdo con los intereses y necesidades de cada estudiante, por lo cual brindar al estudiante esta herramienta le facilita un aprendizaje activo y puede aumentar el interés por los conceptos de su clase.

1.2. Estudios relacionados

En la bibliografía podemos identificar diversos trabajos relacionados al desarrollo de los asistentes conversacionales y a su implementación, son usados para apoyar algún tema específico a las y los estudiantes. De entre los cuales existen: la implementación de un chatbot para apoyar la enseñanza de un curso de algoritmos tal y como lo describen De la Cruz Apari y Mere Jiménez (2023) quienes dentro de su estudio se enfocan en esa materia dado que los mismos estudiantes expresan insatisfacción por no tener información suficiente para entender la materia, motivo por el cual evaluaron el impacto que tiene implementar un chatbot en su curso de algoritmos; concluyendo que generan un cambio significativo en el aprendizaje del estudiantado, con un efecto positivo y respaldan que el uso de su chatbot mejora su aprendizaje.

Bustos Macedo y Romero Loa (2023) estudiaron el impacto que tiene un chatbot en estudiantes universitarios que se dedicará a enseñar acerca del software de control de versiones Git, debido que muchas empresas dedicadas al desarrollo de software utilizan estas tecnologías para su funcionamiento. Los autores detectaron que las y los estudiantes egresados no cuentan con los conocimientos necesarios para utilizar este tipo de servicios, concluyen que al utilizar su chatbot incrementaron la comprensión y habilidades relacionadas con este software de control de versiones.

1.3. Preguntas de Investigación y objetivos

Ante la problemática identificada, este equipo de autores planteamos las siguientes preguntas de investigación: ¿Qué beneficios y opiniones tendrán las y los estudiantes al usar un agente conversacional que les permita resolver preguntas sobre temas de programación de computadoras?, ¿Qué opinión tienen los docentes de la experiencia de interacción con un asistente conversacional dedicado a responder preguntas de programación para ser usado como apoyo en el aprendizaje activo de las y los jóvenes?

Para responder a estas cuestiones se definieron los siguientes objetivos: 1) desarrollar e implementar un agente conversacional que apoye el aprendizaje de programación de computadoras a alumnos que no cuenten con los conocimientos necesarios sobre dicha materia y 2) evaluar la opinión de docentes y alumnos respecto a la funcionalidad, desempeño y resultados de experiencia de usuario al utilizar el agente conversacional.

El público objetivo son los alumnos del primer año de la carrera de Ingeniería en Computación del Centro Universitario UAEM Texcoco, quienes se ven afectados ante los cambios educativos generados a partir de la pandemia del año 2020.

Esto con finalidad de explorar el resultado de usar un chatbot en el sector educacional, pues son herramientas tecnológicas aceptadas por la comunidad estudiantil, que requiere de aplicaciones que apoyen su aprendizaje sobre asignaturas en las que tienen dificultades o en las que no cuentan con los conocimientos previos suficientes para abordar los temas que comprenden las unidades de aprendizaje de los centros educativos.

2. MATERIAL Y MÉTODO

Este estudio se desarrolló en dos fases, la primera corresponde al área de Ingeniería de software, se aplicó el paradigma de desarrollo “espiral” para desarrollar el asistente conversacional. La segunda fase fue un estudio cuantitativo exploratorio sobre los resultados de aprendizaje y opiniones de uso sobre el asistente conversacional. El modelo espiral se realiza en 4 fases: planificación, análisis de riesgo, implementación y evaluación (Ghezzi et al., 2003).

En la planificación se identificaron los requerimientos funcionales del software con base en las características pedagógicas del curso, la población estudiantil y la planta docente. Se definió que el chatbot tiene el objetivo de apoyar el aprendizaje de la POO (Programación Orientada a Objetos), se estableció que la base de conocimientos debe incluir los temas de la asignatura Programación I del programa de estudios de la Licenciatura de Ingeniería en Computación de la Universidad Autónoma del Estado de México (Albarrán Trujillo et al., 2019), específicamente las unidades de aprendizaje uno, dos y tres que incluyen conceptos de:

los elementos de un programa, sintaxis de un lenguaje de POO, su representación en seudocódigo y el Lenguaje Unificado de Modelado (UML por sus siglas en inglés), datos, constantes, variables, clases, objetos, atributos, métodos, parámetros, encapsulamiento, modificadores de acceso, instanciación, diagramas UML, arreglos y estructuras de control.

En esta fase también se exploraron las necesidades e intereses del estudiantado, para ello se recopilaron mediante un cuestionario aquellas preguntas que tentativamente la comunidad estudiantil podría realizarle al agente conversacional. Los resultados también dieron datos sobre la ambigüedad que pudieran tener las y los estudiantes al interactuar con el chatbot. Se acudió a alumnos del primer año, periodo 2023A a quienes se les solicitó que redactaran 20 preguntas que le harían a un experto en programación sobre las dudas que pudieran tener acerca de los temas relacionados con esta asignatura. En la tabla 1 se muestran los resultados obtenidos.

TABLA 1. Levantamiento de preguntas a estudiantes del primer año en el CU UAEM Texcoco

Grupo	Cuestionarios obtenidos	Cantidad Preguntas capturadas
I2	20	387
1I	20	368
2I	22	418
2I1	35	468
Total	**97**	**1641**
Total, preguntas procesadas		**1399**

Fuente: Elaboración propia

Con el banco de preguntas transcritas en un archivo digital de tipo Excel, se realizó el análisis correspondiente del lenguaje y datos lingüísticos que utilizan los estudiantes mediante el uso de la herramienta Lancsbox la cual permite realizar análisis de corpus lingüísticos e identificar aspectos específicos en el lenguaje del estudiantado.

Para realizar el análisis lingüístico en las preguntas que elaboró el estudiantado, se configuró Lancsbox para obtener la frecuencia absoluta de cada una de las palabras contenidas en las preguntas capturadas. A partir de este análisis se identificaron los cinco temas que tuvieron más frecuencia, coincidiendo con los conceptos básicos de la programación, esta información se presenta en la tabla 2.

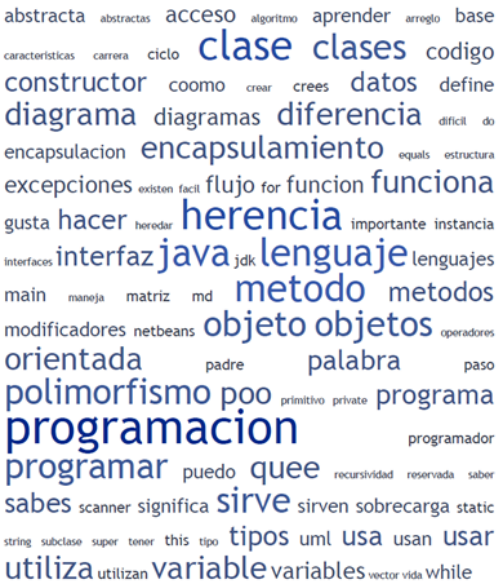
TABLA 2. Temas y preguntas con mayor frecuencia escritas por los estudiantes del CU UAEM Texcoco

Tema académico	Ejemplo de preguntas del estudiantado	Frecuencia por palabra
Herencia	¿Qué es herencia?, ¿Qué es extends?	78
Polimorfismo	No sé qué es polimorfismo, ¿Para qué se usa el polimorfismo?	55
UML	No entiendo el UML, ¿Qué significa UML?	55
Encapsulamiento	¿Qué es encapsulamiento y para qué sirve?	54
Programación Orientada a Objetos	¿Qué es la programación?,¿Para qué se usan las clases?, ¿Qué es una clase?	47
Java	¿Para qué se usa java?, ¿Para qué sirve java?, ¿Por qué se debe de usar java?	22

Fuente: Elaboración propia

Además del análisis lingüístico para identificar los temas en que las y los estudiantes tienen más duda, mediante una nube de palabras se obtuvieron las cien que más se repetían como se muestra en la figura 1.

FIGURE 1. Nube de las cien palabras más frecuentes obtenidas de las preguntas escritas por las y los estudiantes



Fuente: Elaboración propia

Una vez identificados los temas y las palabras que más frecuencia tiene el análisis lingüístico del estudiantado se generaron los grafos no dirigidos de colocación de palabras, mediante la opción de GraphColl en Lancsbox. Los grafos se utilizaron para seleccionar, delimitar y crear el banco de conocimiento junto con el material bibliográfico que recomienda la materia dedicada a la programación del CU UAEM Texcoco. De esta forma se delimitaron los temas que el chatbot abordó.

En el análisis de riesgos se identificaron las fuentes de riesgo para el proyecto, por lo que se hizo un análisis comparativo de las distintas plataformas de desarrollo de chatbots, considerando los siguientes criterios: facilidad de uso, de desarrollo, de integración y accesibilidad para toda la comunidad estudiantil. Se eligió de entre todas a Dialogflow como la plataforma para desarrollar e implementar el chatbot para las y los estudiantes objetivo.

Para minimizar las respuestas superficiales y mostrar respuestas validadas por expertos se definieron las respuestas del chatbot y los flujos de conversación en base a los textos de la bibliografía recomendada como se comentaba anteriormente.

Dentro de la implementación se desarrollaron los flujos de conversación, se hicieron pruebas y se configuraron los canales de integración. En esta fase las primeras pruebas se realizaron en la población docente, se distribuyó el asistente conversacional en un plazo de 2 semanas, el objetivo fue que los docentes evaluaran la pertinencia de las respuestas, que conocieran antes que los alumnos la herramienta. Se aplicaron cuestionarios de evaluación basados en el instrumento de usabilidad de Rauschenberger et al. (2013) y se les pidió que en caso de recibir respuestas incorrectas fueran reportadas. Con los datos obtenidos se decidió liberar el asistente a la comunidad estudiantil para su uso. Para ello se dio por dos semanas y al término del plazo se recopilaron datos de la evaluación de conocimientos y de opinión de usabilidad del estudiantado.

Finalmente, en la evaluación se vinculó con el estudio exploratorio, los detalles se explican en las siguientes secciones del presente artículo.

2.1. Instrumentos de recolección de datos

En el proceso de la investigación se desarrollaron tres instrumentos de recolección de datos:

a. Instrumento de evaluación diagnóstica sobre conocimientos de POO: se seleccionaron 100 reactivos de opción múltiple del banco de reactivos que los docentes formaron colegiadamente durante años anteriores como parte del trabajo de elaboración de exámenes departamentales. Se distribuyó a la población estudiantil en dos momentos: antes y después de usar el asistente conversacional. Para obtener la validación interna los resultados se sometieron a un análisis de coeficiente de correlación mediante la matriz de correlación de los ítems y se obtuvo la tabla 3 que muestra los resultados considerando la escala de interpretación de la magnitud del Coeficiente en Corral (2019).

TABLA 3. Alfa de Cronbach de cada categoría identificada en el instrumento de pretest y post test aplicado a la muestra de alumnos

Categoría de preguntas	Ítems	Alfa de Cronbach	Rango del Coeficiente de correlación	Magnitud
Conceptos de la POO	26	0.73210248	0.70 a 1.00	Muy fuerte
Sintaxis sobre POO con Java	46	0.79791406	0.70 a 1.00	Muy fuerte
Conocimientos sobre arreglos	18	0.77549781	0.70 a 1.00	Muy fuerte
Conocimientos sobre el uso de UML	10	0.56107761	0.50 a 0.69	Sustancial

Fuente: Elaboración propia

b. Instrumento para evaluar la opinión de los docentes sobre el uso del asistente conversacional: para verificar que el asistente fuera una herramienta de apoyo para el aprendizaje del estudiantado se aplicó un cuestionario con preguntas cerradas, abiertas y con escala de Likert, basado en lo documentado por Rauschenberger et al. (2013). La confiabilidad se comprobó mediante el Coeficiente de Alfa de Cronbach, el cual evalúa la confiabilidad u homogeneidad de las preguntas o ítems mediante la varianza de los ítems y la varianza del puntaje total como menciona Corral (2019). Obteniendo resultados de Alfa de Cronbach de 0.95 lo cual indica que el instrumento utilizado es confiable.

c. Instrumento para evaluar la opinión de las y los estudiantes sobre el uso del asistente conversacional: se aplicó un cuestionario con preguntas cerradas, abiertas y con escala de Likert, basado en lo documentado por Rauschenberger et al. (2013). Así mismo para comprobar la confiabilidad del instrumento utilizado para la evaluación del estudiantado se hace uso del Coeficiente Alfa de Cronbach. Teniendo para este instrumento un Alfa de Cronbach de 0.98 siendo un indicador confiable en el instrumento aplicado.

2.2. Muestra

La muestra que se utilizó para obtener la evaluación docente incluye la participación de 10 docentes de los Centros Universitarios Texcoco, Valle de Chalco, Teotihuacán y Nezahualcóyotl. Mientras que, para la

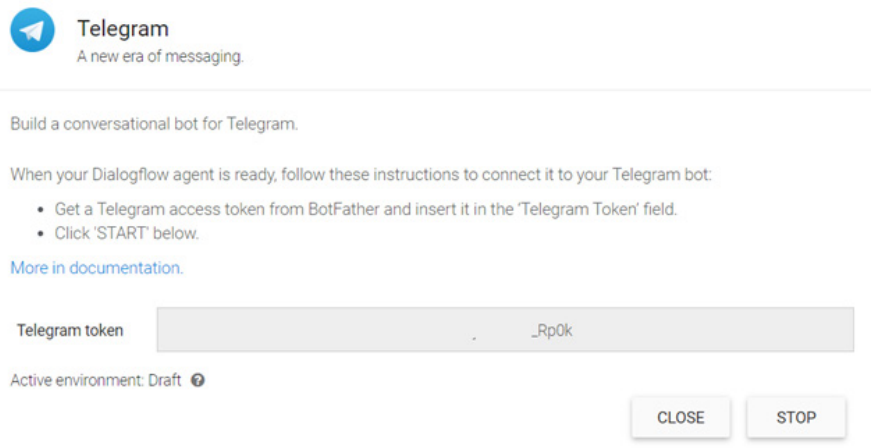
muestra estudiantil, se difundió el chatbot a 4 grupos (90 alumnos) para su uso y para la evaluación se obtuvieron datos de una muestra aleatoria por invitación: siendo un total de 12 alumnos participantes; cabe señalar que hubo más respuestas a la evaluación de uso del chatbot, sin embargo, en este informe sólo se reportan los datos de los cuestionarios que tenían tanto el cuestionario diagnóstico aplicado antes de usar el chatbot, como del que se aplicó después de usar el chatbot.

2.3. Integración y Difusión

Después de registrar un nombre en la aplicación de mensajería Telegram para la vinculación entre la plataforma y el medio de comunicación para centralizar el aprendizaje de los estudiantes; se eligió el nombre de DeviBot para poder difundir el chatbot tanto a docentes como a estudiantes.

La integración entre Dialogflow y Telegram se realiza desde el apartado integraciones en la consola de la plataforma Dialogflow misma que se presenta en la figura 2; en este apartado se añade el token que se genera en Telegram al crear un bot en la aplicación de mensajería instantánea, una vez agregado el token correspondiente se inicia la integración, para que el chatbot tenga como interfaz de usuario la APP de Telegram.

FIGURE 2. Integración entre el agente de Dialogflow y la app de mensajeria instantánea Telegram



Fuente: Elaboración propia

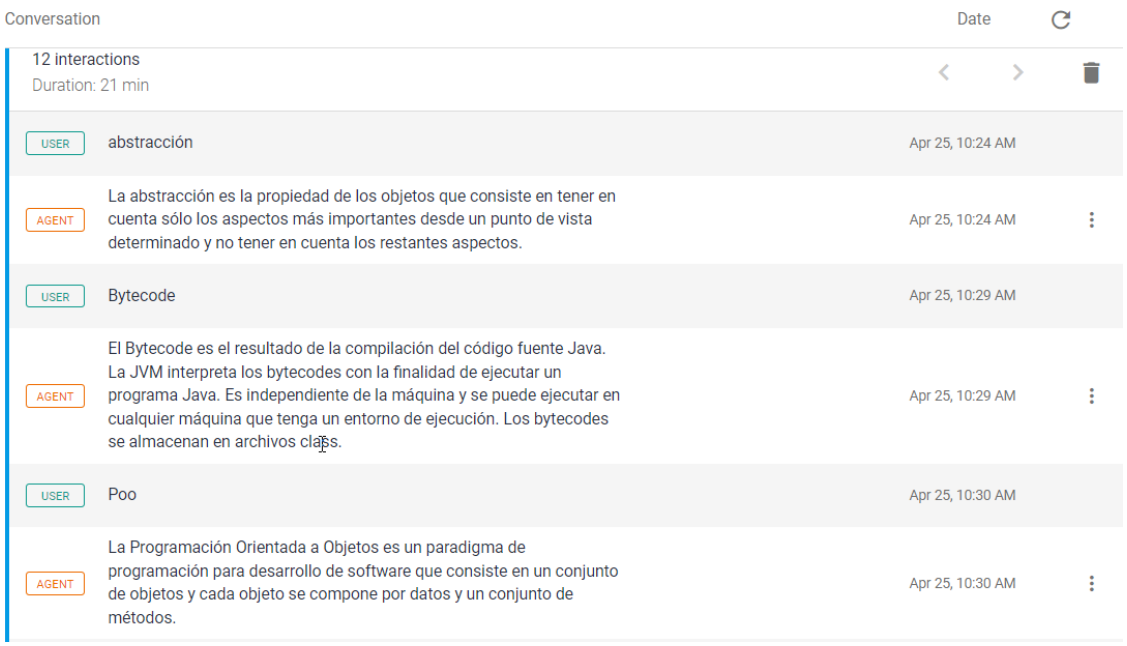
Una vez integrado el agente conversacional de Dialogflow y Telegram en la figura 3 se muestra una interacción utilizando la aplicación de mensajería Telegram y las respuestas presentadas que se obtienen del banco de conocimiento del agente que se encuentran en Dialogflow.

FIGURE 3. Interacción con el agente conversacional de Dialogflow mediante la app de Telegram

Fuente: Elaboración propia

Así mismo Dialogflow permite obtener las conversaciones que se han realizado entre el estudiantado y docentes con el propio agente que ha resuelto las dudas respecto a programación, en la figura 4 se muestra un ejemplo de esta información obtenida desde Dialogflow.

FIGURE 4. Conversación obtenida desde la consola de Dialogflow



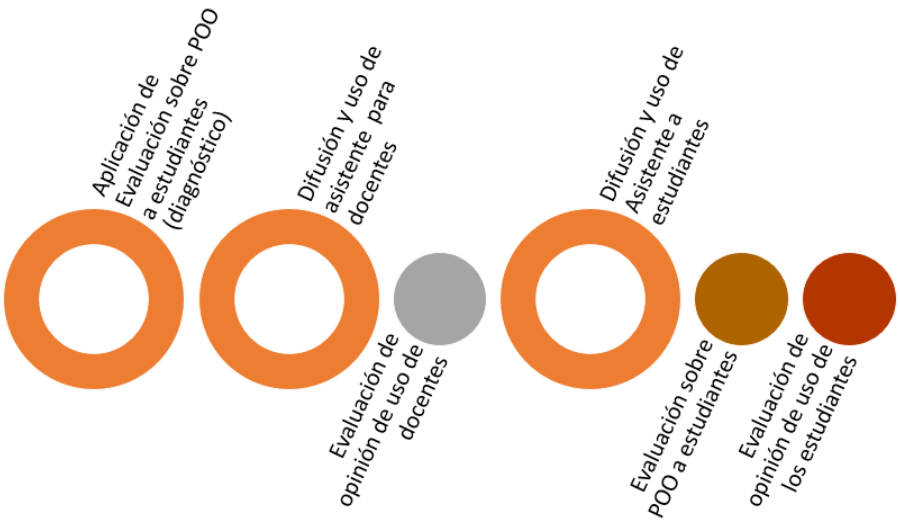
Fuente: Elaboración propia.

De manera tal se confirma que el chatbot está enlazado correctamente para atender las interacciones de los docentes y el estudiantado. Se continuó con la difusión del asistente conversacional a la comunidad y con la ayuda del cuestionario dirigido a los docentes se verificó y validó la información que el agente es capaz de responder, además nos permitió conocer la experiencia de usuario y las posibles mejoras que se pueden añadir al agente para que sea una herramienta útil para la comunidad estudiantil.

Primero se difundió a los docentes (10 personas) en un plazo de dos semanas, al final del periodo establecido se les pidió responder el cuestionario de evaluación y pertinencia, los resultados nos permitieron conocer la experiencia de usuario y las posibles mejoras que se pueden añadir al agente para que sea una herramienta útil para la comunidad estudiantil.

Los datos recabados se consideraron favorables por lo que se difundió a 4 grupos de la Licenciatura en un periodo establecido para que se familiarizaran con la forma de interacción para solventar sus dudas en sus conocimientos generales. Y, por último, se aplicó un cuestionario de experiencia de usuario a los estudiantes para conocer su opinión acerca de la interacción y uso del agente conversacional como herramienta de apoyo en su aprendizaje en programación de computadoras. La figura 5 muestra de forma esquemática los momentos de uso y evaluación del asistente conversacional.

FIGURE 5. Proceso de uso y evaluación del asistente conversacional Devibot



Fuente: Elaboración propia

2.4. Desarrollo de hipótesis

Para evaluar la opinión de los docentes respecto a la experiencia de usuario que tuvieron con el asistente se formularon las siguientes hipótesis:

1. Hipótesis docente 1 (HD1: P ≤ 25%): El porcentaje de profesores que percibe que el comportamiento del chatbot DeviBot tiene características negativas de uso para los alumnos es menor o igual al 25%.
2. Hipótesis docente 2 (HD2 P ≥ 75%): El porcentaje de profesores que percibe que el comportamiento del chatbot DeviBot tiene características positivas de uso para los alumnos es mayor o igual al 75%.

Mientras que para valorar la evaluación de los estudiantes respecto a la experiencia de usuario que tuvieron con el asistente conversacional se realizaron los siguientes planteamientos.

1. Hipótesis estudiantil 1 (HE1: P ≤ 20%): El porcentaje de estudiantes que percibe el comportamiento del chatbot DeviBot con características negativas es menor o igual al 20%.
2. Hipótesis estudiantil 2 (HE2: P ≥ 80%): El porcentaje de estudiantes que percibe el comportamiento del chatbot DeviBot con características positivas es mayor o igual al 80%.

Los porcentajes seleccionados son el 25% (para docentes) o 20% (para los alumnos) que tengan opiniones negativas, pues suponemos que no se superará el primer cuartil de docentes y estudiantes con mala percepción sobre el asistente; mientras que el 75% para docentes y 80% del estudiantado representan los aspectos positivos pues suponemos que más de la media de docentes y estudiantes tendrá buenas opiniones.

3. RESULTADOS

Para medir el grado de concordancia entre los datos recopilados y las hipótesis nulas, aplicamos el estadístico de prueba T de Student, ya que es la prueba adecuada para muestras pequeñas (máximo 30 datos). En este caso se cumple el tamaño de las muestras, tanto para docentes como para la población estudiantil que respondió la evaluación sobre el uso del asistente conversacional. Por lo cual utilizamos la fórmula 1:

FÓRMULA 1. Para calcular la prueba T

Para calcular T de prueba

$$t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

Donde:
t: Valor de T calculada
$\bar{X}$: Media geométrica
μ: Valor por probar
s: desviación estándar muestral
n: tamaño de muestra

Fuente: Obtenida de Hernández Rodríguez et al. (2020)

Para obtener la T tablas para un nivel de confiabilidad aplicaremos la fórmula 2:

FÓRMULA 2. Para obtener la T tablas para un nivel de confiabilidad

Para calcular T tablas

$$DIST.T.INV(2 * \alpha, gl)$$

$$gl = n - 1$$

Donde:
α: Nivel de confiabilidad
gl: Grados de libertad
n: tamaño de muestra

Fuente: Obtenida de Hernández Rodríguez et al. (2020)

Respecto a la opinión de los docentes los datos permiten hacer las pruebas de hipótesis como explican Hernández Rodríguez et al. (2020). En la tabla 4 se muestran los resultados de las pruebas de hipótesis planteadas.

TABLA 4. Resultados por cada prueba de hipótesis de la opinión de los docentes

Escala	Muestra	Grados de Libertad	Nivel de Confiabilidad (α)	Media geométrica	Desviación estándar muestral	Puntaje de acuerdo y totalmente de acuerdo	T tablas Para (α)	Prueba T calculada	Resultado
Experiencia negativa	10	9	97%	19.7665	5.0126	13	2.15	4.2687	Se rechaza HDA1 y se acepta HD1
Experiencia positiva	10	9	97%	36.1964	2.8392	38	-2.15	-2.008	Se acepta HD2

Al interpretar los resultados de las hipótesis para los aspectos negativos con un nivel de confiabilidad al 97% no existe evidencia estadística para afirmar que el porcentaje de profesores que percibe características negativas en el uso del chatbot DeviBot es mayor al 25%; mientras que para los aspectos positivos con un nivel de confiabilidad al 97% existe evidencia estadística para afirmar que el porcentaje de profesores que percibe características positivas en el uso del chatbot DeviBot es mayor al 75%, por lo que se considera que es una herramienta tecnológica de apoyo para los estudiantes que hagan uso de ella.

Para los cálculos respecto a la opinión de los estudiantes los datos obtenidos cumplen con las características de dos muestras en distintos momentos y muestras menores a 30 datos, por lo tanto, se volverán a usar las fórmulas 1 y 2.

En la tabla 5 se muestran los resultados de las pruebas de hipótesis planteadas.

TABLA 5. Resultados de prueba de hipótesis sobre percepción de los alumnos

Escala	Muestra	Grados de Libertad	Nivel de Confiabilidad (α)	Media geométrica	Desviación estándar muestral	Puntaje de acuerdo y totalmente de acuerdo	T tablas Para (α)	Prueba T calculada	Resultado
Experiencia negativa	12	11	95%	20.5310	7.0533	12	1.79	4.189	Se rechaza HEA1 y se acepta HA1
Experiencia positiva	12	11	95%	49.5659	5.1015	48	1.79	1.0633	Se acepta HE2

Al interpretar los resultados de las hipótesis para la evaluación de aspectos negativos con un nivel de confiabilidad del 95% no existe evidencia estadística para afirmar que el porcentaje de alumnos percibe características negativas en el uso del chatbot DeviBot es mayor al 20%; mientras que para los aspectos positivos respecto a la evaluación de los estudiantes con un nivel de confiabilidad del 95% existe evidencia estadística para afirmar que el porcentaje de estudiantes percibe características positivas en el uso del chatbot DeviBot es mayor al 80%.

Así se demostró que el uso del chatbot DeviBot permite a los estudiantes adquirir nuevo conocimiento y aclarar las dudas que tienen sobre la programación de computadoras, además de ser una herramienta de apoyo que el profesor puede utilizar como asesoría personalizada, minimiza el tiempo invertido que el profesor dedica a esta actividad; permitiéndole generar nuevos materiales y ejercicios que son de utilidad para el estudiantado.

4. DISCUSIÓN

El chatbot fue evaluado y probado en primera instancia por docentes para verificar que la información que entrega corresponde al programa de estudios de la unidad de aprendizaje de Programación I de la licenciatura en Ingeniería en Computación de acuerdo con Albarrán Trujillo et al. (2019). Este chatbot fue desarrollado con la plataforma Dialogflow de Google integrado con la aplicación de mensajería instantánea Telegram para aprovechar sus funcionalidades de interacción para con el estudiante y el usuario que lo utilice.

Las pruebas realizadas por los docentes evaluó parte de la ambigüedad que los estudiantes pueden tener al realizarle las preguntas correspondientes al agente conversacional. Considerando que los temas con mayor número de frecuencia por parte del estudiantado son herencia, polimorfismo, UML, encapsulamiento, programación y java respectivamente.

Teniendo las opiniones acerca del funcionamiento sobre el chatbot y retomando los resultados de las pruebas de hipótesis de la evaluación de los docentes se considera que el uso de esta herramienta contiene características positivas de uso en las y los estudiantes ya que es mayor al 75%, por lo que los docentes pueden invertir el tiempo ahorrado por el chatbot al dar las asesorías para otras actividades enfocadas en las y los estudiantes.

Así mismo con la afirmación de García Brustenga et al. (2018) sobre los agentes conversacionales acerca de que aportan personalización de aprendizaje, mejorando la metacognición y aumentando la motivación en los estudiantes, se sustentan los objetivos de esta investigación la cual consistió en desarrollar e implementar un agente conversacional que apoye el aprendizaje de programación de computadoras.

Con los resultados obtenidos de la evaluación de las y los estudiantes mediante la prueba estadística correspondiente se confirma que el 80% de ellos percibe características positivas al utilizar el agente conversacional y se concluye que esta herramienta tecnológica ayuda a su aprendizaje autodirigido.

Partiendo de que el estudiantado puede adquirir conocimiento nuevo y aclarar sus dudas acerca de programación con el uso de un agente conversacional especializado y focalizado que responde a sus preguntas para maximizar las asesorías personalizadas, minimizando el tiempo que debe invertir el docente a esta actividad en específico e invertirlo en otras actividades que permiten ser de utilidad al estudiantado para mejorar su enseñanza aprendizaje acerca de una asignatura en específico, donde se ha identificado que nuevas herramientas aplicadas a la educación permiten apoyar el aprendizaje de las y los estudiantes y motivarlos a consultar fuentes de información especializada.

Mohamed et al. (2024) comentan que la Inteligencia Artificial permite mejorar la experiencia del aprendizaje de las y los estudiantes, ya que proporcionan métodos creativos para difundir conocimiento además de que motiva a los estudiantes para aprender dado el nivel de compromiso con su aprendizaje.

El uso de estas herramientas tecnológicas donde se incluye Inteligencia Artificial motiva al estudiantado ya que permite mejorar los resultados mediante la aplicación de estrategias de enseñanza directamente, demostrando que existe una relación inherente entre motivación y resultados.

Al aumentar la motivación de los estudiantes por aprender, se mejora su rendimiento escolar haciendo que los docentes se enfoquen en materiales educativos interactivos para mejorar el interés del estudiantado para utilizar herramientas tecnológicas como pueden ser los chatbots aplicados en el ámbito escolar.

Además, el entorno educativo se puede ver beneficiado por el aprendizaje personalizado al hacer uso de la Inteligencia Artificial, siendo los sistemas de tutoría un aliciente para mejorar el rendimiento académico de las y los estudiantes. Por lo que los agentes conversacionales han venido a cambiar la educación tradicional tal como lo comenta Mohamed et al. (2024).

Ahora bien, respecto a lo que comentan Principal y Orellana (2024) el uso de chatbots educativos como aplicación de Inteligencia Artificial permite una tutorización virtual en tiempo real, en donde se obtienen respuestas a dudas, ofreciendo explicaciones y sirven de guía al estudiantado a través de actividades que aporten en su enseñanza aprendizaje, además comentan que esta estrategia de aprendizaje ahorra tiempo y elimina distracciones por tener conversaciones largas, dado que se gestiona el propio aprendizaje con base a la información que se solicite entre el chatbot y el estudiantado.

5. CONCLUSIONES

El trabajo realizado se enfoca en documentar cómo un agente conversacional permite ser una herramienta especializada para apoyar la enseñanza y aprendizaje sobre temas en específico. El uso de los chatbots en el ámbito educativo es aceptado por la comunidad estudiantil ya que requieren de herramientas que los apoyen en su aprendizaje de aquellas materias que tienen dificultades o en las que sus conocimientos son limitados. El estudio permitió determinar que estas aplicaciones tecnológicas apoyan a los y las estudiantes a incrementar sus conocimientos y habilidades.

Si bien la Inteligencia Artificial se considera cada vez más poderosa que permite aumentar las capacidades de pensamiento crítico en el entorno educativo, siendo los chatbots plataformas de aprendizaje autodirigido y autorregulado permitiendo ser asistentes de aprendizaje personalizado como comentan Darwin et al. (2024).

Los chatbots abren un debate sobre los métodos de enseñanza de las instituciones educativas ya que pueden ser de apoyo para el aprendizaje de las y los estudiantes tanto como una distracción, ya que surge la incógnita de que si realmente los y las jóvenes están aprendiendo sobre aquello que necesitan, sin embargo, Tiwari (2024) menciona que el aprendizaje personalizado representa una atención histórica a los planes de estudio, ya que dicho estilo de aprendizaje ahora permite comprender como aprende el estudiantado y ajustar los planes de estudio respecto a sus necesidades.

Es por ello por lo que los chatbots se asemejan a los Sistemas de Tutoría Inteligente ya que funcionan sin necesidad de tener un profesor presente y se utilizan diversos algoritmos en ellos para retar y ayudar al estudiantado de forma eficaz, además de guiar con ejemplos dado que se orienta a las y los estudiantes a identificar técnicas eficaces para resolver problemas utilizando modelos de trabajo en equipo para cuando se requiera. Siendo la principal ventaja del uso de chatbots en el ámbito educativo el aprendizaje adaptativo diseñando una experiencia educativa que guie las competencias del estudiantado en una enseñanza simultánea, es decir dentro y fuera del aula de clase.

Los chatbots permiten a las y los estudiantes tener una herramienta como fuente de consulta de información disponible cuando sus profesores de asignatura no cuenten con la disponibilidad de horario y lugar para realizar una asesoría presencial, además como apoyo a los docentes les permite minimizar tareas repetitivas como es la explicación de los mismos temas a distintos estudiantes en tiempos de disponibilidad distintas entre uno y otro; donde pueden ocupar ese tiempo en diseñar nuevos ejercicios y materiales didácticos para agregarlos dentro de los temas que contiene el chatbot para que el estudiantado pueda comprender mejor un tema por tener más información especializada y focalizada.

Si bien quedan abiertas varias preguntas de investigación para comprender las implicaciones y el nivel de apropiación del aprendizaje logrado al utilizar herramientas como son los agentes conversacionales Shaikh et al. (2023) hacen un llamado de acción a distintos actores de entre los cuales rescatamos a los académicos para investigar las ventajas, desventajas y sesgos que pueden tener en el estudiantado la aplicación de este tipo de herramientas tecnológicas para apoyar su aprendizaje, en donde además incluyen a los diferentes lideres del sector para adoptar practicas responsables tanto en desarrollo como en aplicación de la Inteligencia Artificial.

De esta forma para nuestra investigación se confirma que el uso de herramientas como son los chatbots funcionan para la comunidad estudiantil del Centro Universitario UAEM Texcoco como tutores personalizados disponibles las 24 horas los 7 días de la semana, únicamente se requiere tener un dispositivo con la aplicación de mensajería instalada y conexión a internet para utilizarlo. Retomando los resultados estadísticos tanto de los docentes como de las y los estudiantes se confirma que estas herramientas sirven para atraer la atención y aumentar el rendimiento del estudiantado al utilizar un agente conversacional en su proceso de aprendizaje.

5.1. Limitaciones y prospectiva

Para incrementar las capacidades del agente conversacional es conveniente incluir temas más complejos de programación y generar materiales didácticos para añadirse a las respuestas que provee el chatbot de modo tal que el estudiantado tenga la suficiente información que le ayude a incrementar su conocimiento y por ende mejorar su rendimiento escolar.

Dentro de esos materiales didácticos se pueden considerar ejemplos, ejercicios, lecturas y cualquier material que sea provechoso para la enseñanza y aprendizaje de los estudiantes, considerando que los docentes serán aquellos que diseñen y realicen dichos materiales dado que el chatbot permite optimizar sus tiempos y reusar los materiales generados en atención a sus estudiantes.

6. REFERENCIAS

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*, *2*, 100006. https://doi.org/10.1016/J.MLWA.2020.100006

Albarrán Trujillo, S. E., Trujillo Flores, E., Vera Noguez, S., Granda Gutiérrez, E. E., y Aguilar Juárez, I. (2019). *Programa de Estudios - Programación I.* Universidad autónoma del estado de México. http://ri.uaemex.mx/bitstream/handle/20.500.11799/107801/secme-12408_1.pdf?sequence=1

Arredondo Castillo, C. C. (2021). *Inteligencia artificial en la educación: uso del chatbot en un curso de pregrado sobre Investigación Académica en una universidad privada de Lima*. [Tesis de doctorado, Pontificia Universidad Católica del Perú]. PUCP http://hdl.handle.net/20.500.12404/20996

Astruga, J. A. (2021). *Propuesta metodológica para el análisis y diseño de chatbots basados en texto* [Trabajo de Fin de Máster, Universidad de Valladolid]. UVaDOC. https://uvadoc.uva.es/handle/10324/50064

Bagchi, M. (2020). Conceptualising a Library Chatbot using Open-Source Conversational Artificial Intelligence. *DESIDOC Journal of Library & Information Technology*, *40*(6), 329–333. https://doi.org/10.14429/djlit.40.06.15611

Berumen López, E., Villegas Berumen, H. G., y Ávila Acevedo, S. (2023). Implicaciones de la educación virtual durante la pandemia covid-19: una encuesta a estudiantes del Tecnológico Nacional de México. *RIDE. Revista Iberoamericana Para La Investigación y El Desarrollo Educativo, 13*(26), e013. https://doi.org/10.23913/RIDE.V13I26.1404

Bustos Macedo, C. J., y Romero Loa, J. A. (2023). *Chatbot para el aprendizaje de la herramienta Git* [Tesis doctoral, Universidad César Vallejo]. Repositorio Institucional - UCV. https://repositorio.ucv.edu.pe/handle/20.500.12692/133458

Carrazco Zendejas, S. G. (2021). La problemática educativa en México: lo que dejó la pandemia por COVID-19. *Dilemas Contemporáneos: Educación, Política y Valores, 8*(23), 1–19. https://doi.org/10.46377/dilemas.v8i.2775

Cervantes H. R. & Chaparro M. P.M. (2021), Transformations in communication and sociability habits through the increased use of socio-digital networks in times of pandemic, *Ámbitos: revista internacional de comunicación, 52*, 37-51. http://doi.org/10.12795/Ambitos.2021.i52.03

Corral, Y. (2019). Validez y Confiabilidad de los instrumentos para la recolección de datos. *Revista Ciencias de La Educación*, *19*(33), 228–247.

Cujia Berrío, S. E. (2023). Transfiguraciones educativas en contextos de pandemia. Educación virtual y presencial en conflicto. *Revista Venezolana de Gerencia, 28*(102), 650–664. https://doi.org/10.52080/rvgluz.28.102.13

Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, *11*(1), 1–18. https://doi.org/10.1080/2331186X.2023.2290342

De la Cruz Apari, J. K., y Mere Jiménez, A. W. (2023). *Implementación de un chatbot con Inteligencia Artificial para mejorar el aprendizaje de algoritmos en los estudiantes de una universidad pública* [Tesis Doctoral, Universidad Autónoma de ICA]. Repositorio Institucional Universidad Autónoma de ICA. https://hdl.handle.net/20.500.14441/2682

Deshpande, B., & Chandak, M. B. (2022). A survey of designing tools for chatbot applications. *International Journal of Health Sciences*, *6*(S5), 1403-1413. https://doi.org/10.53730/ijhs.v6nS5.8889

García Brustenga, G., Fuertes Alpiste, M., & Molas Castells, N. (2018). Briefing paper: chatbots in education. *Barcelona: ELearn Center*, 1–33. https://doi.org/10.7238/elc.chatbots.2018

García Martínez, V., y Silva Payró, M. P. (2022). Percepción académica sobre las barreras en la adopción de innovaciones tecnológicas durante la pandemia por la covid-19. *Apertura, 14*(1), 96–113. https://doi.org/10.32870/AP.V14N1.2150

Ghezzi, C., Jazayeri, M., & Mandiorli, D. (2003). *Fundamentals of Software Engineering* (Second Edition). Prentice Hall.

Hernández Rodríguez, R., Cárdenas Arriaga, T. N., y Hernández Rodríguez, N. A. (2020). *Prueba de Hipótesis Estadística con Excel*. Amat Editorial.

Maher, S., Kayte, S., & Nimbhore, S. (2020). Chatbots & Its Techniques Using AI: a Review. *International Journal for Research in Applied Science and Engineering Technology, 8*(XII), 503-508. https://doi.org/10.22214/ijraset.2020.32537

Mendoza Navarro, F. B. (2022). La Tutoría Virtual: Un Soporte Efectivo para la Estrategia Aprendo en Casa Durante la Emergencia Educativa del Covid-19. *Revista Científica de Salud y Desarrollo Humano, 3*(1), 60–72. https://doi.org/10.61368/r.s.d.h.v3i1.44

Mohamed, A. M., Shaaban, T. S., Bakry, S. H., Guillén Gámez, F. D., & Strzelecki, A. (2024). Empowering the Faculty of Education

Students: Applying AI's Potential for Motivating and Enhancing Learning. *Innovative Higher Education*, 1–23. https://doi.org/10.1007/S10755-024-09747-Z

Morales Bonilla, Y., y Bustamante Peralta, K. E. (2021). Retos de la enseñanza en la pandemia por COVID 19 en México. *Dilemas Contemporáneos: Educación, Política y Valores, 8*(21), 1–18. https://doi.org/10.46377/dilemas.v8i.2772

Ortega Sánchez, R. M. (2021). Uso de Herramientas Tecnológicas en Tiempos de COVID-19. *Revista Docentes 2.0, 12*(1), 31–39. https://doi.org/10.37843/rted.v1i1.223

Ponce Ceballos, S., Martínez Soto, Y., Ruelas Mexía, P. Y., y Toledo Sarracino, D. G. (2022). Tutorías Académicas durante la contingencia Académica por la Covid-19: La óptica de estudiantes de educación superior. *Education Policy Analysis Archives, 30*(17), 1–26. https://doi.org/10.14507/epaa.30.6852

Principal, M., y Orellana, A. (2024). Inteligencia artificial en la enseñanza-aprendizaje universitaria: una revisión semi-sistemática. *EVSOS - Educación y Vida Sostenible*, *2*(4), 23–46. https://doi.org/10.57175/EVSOS.V2I4.157

Ramos Mendoza, J. C., Tasayco Jala, A. A., Pachas Barrientos, L. M., y Valdez Arango, A. (2023). Enseñanza remota como viabilizador de la educación superior pública en tiempos de pandemia. Horizontes. *Revista de Investigación En Ciencias de La Educación, 7*(30), 1804–1820. https://doi.org/10.33996/revistahorizontes.v7i30.630

Rauschenberger, M., Schrepp, M., Perez-Cota, M., Olschner, S., & Thomaschewski, J. (2013). Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ). Example: Spanish Language Version. *International Journal of Interactive Multimedia and Artificial Intelligence*, *2*(1), 39. https://doi.org/10.9781/IJIMAI.2013.215

Shaikh, S., Mhaske, S., Bendre, R., & Aggarwal, A. (2023). The Rise of Creative Machines: Exploring the Impact of Generative AI. *ArXiv*: 2311.13262v1. https://doi.org/10.48550/arXiv.2311.13262

Thorat, S. A., & Jadhav, V. D. (2020). A Review on Implementation Issues of Rule-based Chatbot Systems. *Proceedings of the International Conference on Innovative Computing and Communication (ICICC, 2020)*. http://doi.org/10.2139/ssrn.3567047

Tiwari, E. S. (2024). *Harnessing the Power of Artificial Intelligence in Education*. http://doi.org/10.2139/ssrn.4747352

Varela Tapia, E. A., Sangacha Tapia, L. M., Acosta Guzmán, I. L., y Celi Párraga, R. J. (2022). Inteligencia Artificial Conversacional para el diseño de un asistente virtual interactivo. *Memorias de La Décima Segunda Conferencia Iberoamericana de Complejidad, Informática y Cibernética (CICIC 2022)*. https://doi.org/10.54808/CICIC2022.01.237

Villón Cabrera, N. (2020). *Inteligencia Artificial aplicada al marketing: Impacto del uso de Chatbots Cognitivos en la satisfacción del cliente del sector bancario* [Trabajo de Fin de Grado, Universidad Peruana de Ciencias Aplicadas]. Repositorio Académico UPC. https://repositorioacademico.upc.edu.pe/handle/10757/652700

Vishwakarma, A., & Pandey, A. (2021). A Review & Comparative Analysis on Various Chatbots Design. *International Journal of Computer Science and Mobile Computing*, *10*(2), 72–78. https://doi.org/10.47760/ijcsmc.2021.v10i02.011

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 92-107 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20372

Impact of a Mobile Application to Improve the Reading and Writing Skills of Immigrant Students in Primary Education

Impacto de una aplicación móvil para mejorar la competencia lectoescritora del estudiante migrante en la etapa de Educación Primaria

RECEIVED 30/07/2024 ACCEPTED 26/09/2024 PUBLISHED 01/06/2025

Inmaculada Aznar Díaz
ORCID 0000-0002-0018-1150
Universidad de Granada, España
iaznar@ugr.es

José María Romero Rodríguez
ORCID 0000-0002-9284-8919
Universidad de Granada, España
romejo@ugr.es

Fernando Lara Lara
ORCID 0000-0003-1545-9132
Universidad de Granada, España
fernandolara@ugr.es

Blanca Berral Ortiz
ORCID 0000-0001-8139-8468
Universidad de Granada, España
blancaberral@ugr.es

ABSTRACT

This work addresses the impact of a mobile application designed to improve the literacy skills of migrant students in Primary Education in Spain. Since the 1990s, Spain has experienced a significant increase in migrants, leading to an educational environment characterized by cultural and linguistic diversity. Literacy, defined not only as the ability to read and write but also as a crucial skill for social and professional integration, becomes a key focus of study (Toledo-Vega et al., 2022). The research adopted a quasi-experimental design with control and experimental groups, selecting 44 students aged 6 to 7 years through convenience sampling, who were randomly assigned to the groups. The Diagnostic Scale for Assessing Literacy Competence (EDECL) from the Junta de Andalucía was used as the data collection instrument. Participants were evaluated using pre-tests and post-tests to measure the impact of the intervention. The results of the statistical analysis showed that the experimental group, which used the mobile application AppRende, demonstrated significant improvements in their literacy competence, increasing their average from 6.33 to 8.75, with a p-value < .001. In contrast, the control group, which used printed materials, showed no significant improvements. These findings indicate that the use of technological tools may be more effective than traditional methods in teaching literacy to migrant students, suggesting the need to incorporate mobile technologies in diverse educational contexts to support their integration and academic success.

KEYWORDS Mobile application; literacy competence; migrant students; technologies; Primary Education.

RESUMEN

Este trabajo aborda el impacto de una aplicación móvil diseñada para mejorar las competencias de alfabetización de los estudiantes migrantes en la Educación Primaria en España. Desde la década de 1990, España ha experimentado un incremento significativo de migrantes, lo que ha generado un entorno educativo caracterizado por la diversidad cultural y lingüística. La alfabetización, definida no solo como la capacidad de leer y escribir, sino también como una habilidad crucial para la integración social y profesional, se convierte en un foco esencial de estudio (Toledo-Vega et al., 2022). La investigación adoptó un diseño cuasi-experimental con grupos de control y experimental, seleccionando a 44 estudiantes de entre 6 y 7 años mediante un muestreo por conveniencia, quienes fueron asignados aleatoriamente a los grupos. Se utilizó la Escala Diagnóstica para Evaluar la Competencia en Alfabetización (EDECL) de la Junta de Andalucía como instrumento de recolección de datos. Los participantes fueron evaluados mediante pre-test y post-test para medir el impacto de la intervención. Los resultados del análisis estadístico mostraron que el grupo experimental, que utilizó la aplicación móvil AppRende, presentó mejoras significativas en su competencia en alfabetización, aumentando su media de 6.33 a 8.75, con un valor de $p < .001$. En contraste, el grupo de control, que utilizó materiales impresos, no mostró mejoras significativas. Estos hallazgos indican que el uso de herramientas tecnológicas puede ser más efectivo que los métodos tradicionales en la enseñanza de la alfabetización a estudiantes migrantes, sugiriendo la necesidad de incorporar tecnologías móviles en contextos educativos diversos para apoyar su integración y éxito académico.

PALABRAS CLAVE Aplicación móvil; competencia lectoescritora; estudiantes migrantes; tecnologías; Educación Primaria.

1. INTRODUCTION

Since the 1990s, Spain has received a significant influx of migrants from Latin America, other European Union countries, and North Africa (González-Monteagudo & León-Sánchez, 2020). As a result, the current educational landscape is characterised by notable complexity due to cultural, linguistic, and migratory factors.

One of the key issues requiring attention is the literacy of the migrant population in Primary Education (PE) classrooms (Abad-Quintanal, 2022; Calvo, 2019). Literacy is a crucial component for social and professional integration, encompassing not only the ability to read and write but also skills that enable individuals to navigate daily life, facilitate access to information, and promote a love of reading and writing (Toledo-Vega et al., 2022).

Recent studies show that migrant students in the Primary Education stage tend to exhibit lower academic performance compared to their native peers (Escarbajal et al., 2019; Rodríguez et al., 2020; Tajic & Bunar, 2023). Additionally, they experience difficulties in translating and interpreting written texts, suffer from anxiety and confusion in the school environment (Ceballos-Vacas & Trujillo-González, 2021; El Jouhri et al., 2022), and have higher rates of failure and early school dropout (Bayón-Calvo, 2019). The exclusion of these students from school environments and their low participation in extracurricular activities highlight the lack of adequate integration mechanisms. The social inclusion of these students is directly linked to their cultural proximity and language proficiency. Furthermore, it is important to emphasize that a lack of linguistic competence can lead to absenteeism, insufficient information, and conflicts within the educational environment (Abad-Quintanal, 2022).

Regarding methods for teaching literacy to migrant pupils, classic authors such as Ferreiro and Teberosky (1979) advocate for a combination of synthetic and analytic approaches, which allows teaching to be tailored to the diverse needs of students, based on up-to-date literature and proven methods. Moreover, studies like those by Calvo (2019) emphasise the importance of creating shared reading spaces that foster a written culture in welcoming environments, while El Jouhri et al. (2022) demonstrate that the presence of trained staff reduces anxiety levels among migrant students.

In this context, mobile technology, including mobile learning tools, stands out as a significant resource for promoting digital inclusion and facilitating access to literacy education (Rosell-Aguilar, 2017).

The purpose of this study is to evaluate a custom-designed mobile application aimed at improving the literacy skills of migrant students in the Primary Education stage. The importance of this research lies in identifying new learning opportunities through the use of mobile technology, which allows for personalised tracking of student progress. The application incorporates various virtual worlds, each focused on a specific dimension of literacy competence, as measured in the pre-test and post-test. These dimensions include reading comprehension, written production, vocabulary, and reading fluency, allowing for a comprehensive and individualized assessment of each student's progress. This study aims to provide empirical evidence on the usefulness of mobile technology as an educational intervention tool.

2. LITERATURE REVIEW

2.1. Definition of mobile learning

Mobile learning, or m-learning, has been defined as the use of mobile devices to generate knowledge and develop skills in a flexible and effective manner (Rosell-Aguilar, 2017). This approach enables students to engage in learning activities that adapt to their pace and lifestyle, promoting greater autonomy in the process of acquiring competencies.

The UNESCO report (2022) highlights the importance of mobile learning in the global educational context, particularly in areas with low connectivity or limited resources. In this regard, mobile devices serve as facilitators of information access, supporting the continuation of learning beyond the classroom and the educational inclusion of students with specific needs, such as migrant students. Furthermore, Bai (2019) argues that mobile learning allows migrant students to interact with peers and educational content at their own pace, providing them with an opportunity for integration into both the educational and social environment.

2.2. Use of mobile applications in education

The use of mobile applications has been widely researched in the educational field in recent years, with numerous studies highlighting their ability to promote ubiquitous, personalized, and autonomous learning. These tools allow students to access educational resources anytime and anywhere, facilitating learning activities without temporal or spatial constraints (Rosell-Aguilar, 2017). This approach not only increases educational opportunities but also promotes digital inclusion for groups facing barriers to formal education.

Various studies have demonstrated the positive impact of mobile applications on the development of literacy skills. For example, Jiménez-García and Martínez-Ortega (2017) analyzed the use of "Legendi," a mobile app designed to teach literacy to illiterate individuals using Android tablets. The results indicated that this application not only facilitated the teaching-learning process but was also effective and user-friendly, significantly improving reading and writing skills.

More recently, research by Máñez-Carvajal and Cervera-Mérida (2022) focused on mobile applications designed for children with specific difficulties in acquiring reading and writing skills. Their findings confirmed that the use of technology-based mobile tools can enhance these competencies. Additionally, Fernández-Batanero et al. (2022) emphasized the value of interactive resources such as e-books and digital games, which not only improve reading and writing skills but also foster a playful and motivating learning experience, granting greater autonomy to students.

The potential of mobile applications is not limited to cognitive development but also contributes to creating more inclusive learning environments. According to Criollo et al. (2021) and Hinojo-Lucena et al. (2020), these tools enable the personalization of educational experiences, adapting to the specific needs of each student, which is particularly crucial in multicultural contexts where linguistic and cultural diversity must be considered.

These tools have also proven effective in developing literacy skills among students with hearing impairments (Martínez-Acosta et al., 2022) and migrant students learning Spanish as a second language (Cáceres-Reche et al., 2024). Other studies, such as those by Hernández-Campos et al. (2020) and García-Rodríguez and Gómez-Díaz (2016), have shown how mobile applications can facilitate literacy learning in students studying Spanish as a foreign language, highlighting their potential to improve academic performance and school integration.

In another study, Hautala et al. (2020) analyzed reading difficulties among first- to fourth-grade primary students using a tablet-based computerized system. When comparing the benefits of this approach to traditional paper-and-pencil assessments, the results showed that using interactive games was effective in identifying students with reading difficulties, particularly in terms of fluency, and to a lesser extent, accuracy. Furthermore, this system supported the gradual development of reading skills across several educational levels.

While the use of mobile technologies is crucial for improving literacy skills among migrant students, it is also essential to address teacher training in interculturality and linguistic sensitivity. These factors not only improve social inclusion but also enrich the educational environment overall (Biasutti et al., 2021). Literacy, as noted by Acedo-Tapia and Maqueda-Berrocal (2022), encompasses not only reading and writing skills but also critical and reflective communication competencies that allow students to develop holistically in their social and professional environments.

However, despite the growing interest in mobile technology for education, further research is needed to explore how these tools can be applied effectively in multicultural and multilingual contexts. In this regard, the Horizon Report (EDUCASE, 2022) highlights the expansion of these technologies in the coming years as crucial tools for improving literacy competence in diverse educational settings.

Despite existing research on appropriate teaching methodologies, there remains a gap in educational interventions specifically addressing the needs of migrant students in the early stages of Primary Education, and the effectiveness of such interventions is still under question.

Based on these considerations, the purpose of this work is to determine the impact of a custom-designed mobile application aimed at improving the literacy skills of migrant students in the Primary Education stage. To this end, the following hypotheses were proposed:

H1: There are significant differences in literacy skills between student groups based on gender and age.

H2: Students who engage in literacy activities through a mobile application achieve significantly greater literacy skills compared to the control group.

3. METHODOLOGY

A quasi-experimental design with control and experimental groups and pre-test and post-test measures was adopted. The sample selection procedure was by convenience. However, the assignment of the treatment to the student groups was carried out randomly.

3.1. Participants

The composition of the groups included a total of 44 students aged between 6 and 7 years. The control group consisted of 20 students (60% boys and 40% girls), while the experimental group comprised 24 students (66.6% boys and 33.4% girls). In this study, as noted by Chou and Feng (2019), it is important to highlight that sample size was not an influencing factor when making comparisons within a single population of students.

The participants were migrant students from primary schools in the northern area of Granada. The sample selection was based on convenience and accessibility, according to a series of criteria: (i) first-year Primary Education students who are learning to read and write; (ii) students with migrant status, whose mother tongue is not Spanish.

The invitation to participate in the study was made personally and conducted outside of school hours. All procedures performed adhered to the ethical standards of the institutional research committee and the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Accordingly, the legal guardians of the underage students were informed of the research purpose, the anonymous treatment of their data, and they gave their informed consent before participating in the study. Additionally, the research received approval from the Ethics Committee of the University of Granada (registration number: 3493/CEIH/2023).

Finally, the age and gender of the students were identified, and a diagnostic scale was used to assess their literacy skills. The sample consisted of 28 boys and 16 girls, aged between 6 and 7 years (M = 6.36; SD = .48).

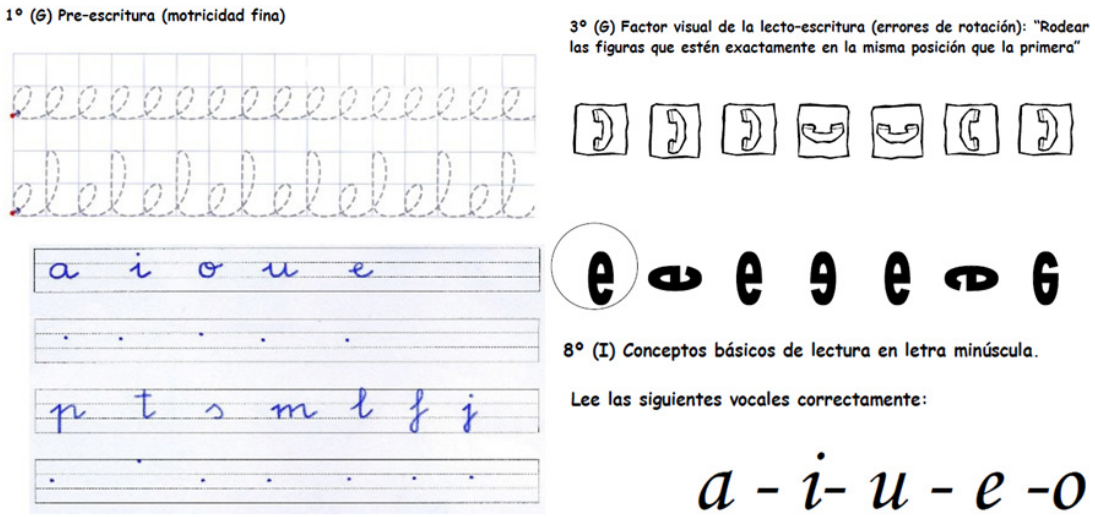
The pre-test was administered on 23 May 2023, while the post-test was administered on 7 June 2023.

3.2. Instrument

The Diagnostic Scale for Assessing Literacy Competence (EDECL) of the Junta de Andalucía (2011) was used. The EDECL was developed considering the concepts of Basic Competences established in the educational regulations of the Autonomous Community of Andalusia (LEA, Ley de Educación de Andalucía).

This instrument measures literacy competence through eight items with various reading and writing comprehension tests. Each item is linked to a specific factor: 1. Pre-writing (fine motor skills); 2. Identifying and writing their names in uppercase letters; 3. Visual factor of literacy (rotation errors); 4. Lateralised Body Schema (in graphic space); 5. Visual Recognition Memory; 6. Auditory Memory; 7. Basic concepts of writing in lowercase letters; 8. Basic concepts of reading in lowercase letters (Figure 1).

FIGURE 1. Examples of items



Regarding the scores, the questionnaire uses a scale from 0 to 10, where 10 represents the highest score in literacy competence. Each item has its specific method of correction; the following are the specific rules for the application and correction of the literacy test:

- **Item 1.** Pre-writing (Maximum 1 point): 1 point if the tracing is perfect; 0 points if the tracing is not perfect.
- **Item 2.** Identification and writing of their name in uppercase (Maximum 1 point): 1 point if they write their name correctly and with all letters; 0 points if letters are missing.
- **Item 3.** Visual factor of literacy (Maximum 1 point): 1 point if all figures are correct; 0.5 points if 2 figures are correct; 0 points if all figures are circled.
- **Item 4.** Lateralised body schema in graphic space (Maximum 1 point): 1 point if both sequences are correct; 0 points if 1 or both sequences are incorrect.
- **Item 5.** Visual recognition memory (Maximum 1 point): 1 point if 3 or 4 drawings are correctly circled; 0.5 points if 1 or 2 drawings are correctly circled; 0 points if all drawings are circled.

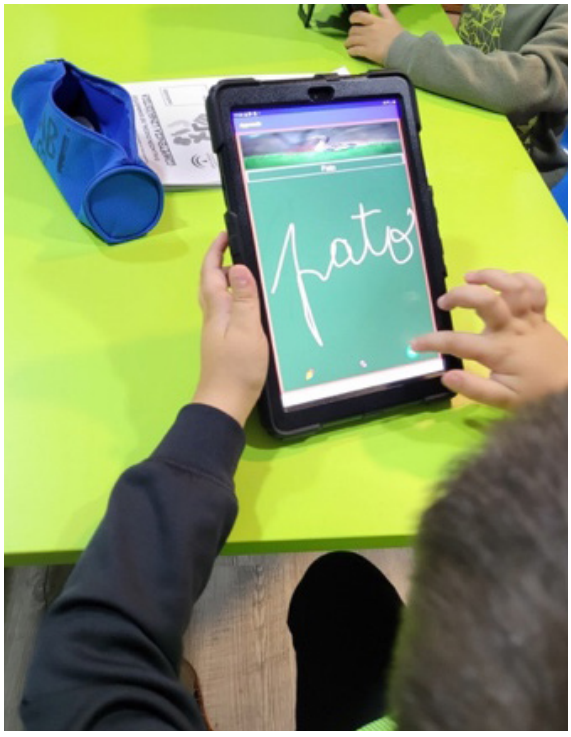
- **Item 6.** Auditory memory (Maximum 1 point): 1 point if all 3 series are correct; 0 points if not all series are correct.
- **Item 7.** Basic concepts of writing. Dictation of vowels (Maximum 2 points): 2 points if all vowels are correct; 1 point if there is 1 mistake; 0 points if there is more than 1 mistake.
- **Item 8.** Basic concepts of reading. Identifying vowels (Maximum 2 points): 2 points if all vowels are read correctly; 0 points if not all vowels are read correctly.

In addition to being a widely accepted and reliable instrument utilized in research by the Junta de Andalucía, the scale demonstrated strong internal consistency in this study. The reliability analysis, conducted using Cronbach's alpha, indicated satisfactory values. Specifically, the pre-test yielded a reliability coefficient of $\alpha = .78$, while the post-test showed an even higher reliability of $\alpha = .85$. These values suggest that the scale provides a robust measure of literacy competence, ensuring consistency in the responses across both testing phases. Such reliability supports the validity of the instrument in evaluating literacy skills in young migrant students and reinforces its utility for future research in similar educational contexts.

3.3. Procedure and materials

In the experimental group, composed of twenty-four students, one teacher was assigned to twelve students, while the other teacher managed the remaining students, forming groups for the corresponding intervention. The experimental group implemented a digital intervention program to enhance reading and writing skills. The program spanned approximately 12 hours, distributed over eight sessions of 1.5 hours each within two weeks. The methodological approach was active and participative, utilising a mobile application called AppRende (Figure 2). This specific application was chosen for its customised design and development for the study, aimed at exploring and evaluating its impact in a specific and detailed manner. Each student had a personal tablet equipped with the pre-installed application; the device used was a Samsung Galaxy Tab A (2019) with the Android 10 operating system. The download link for the application is available upon request.

FIGURE 2. Student in the intervention phase



The designed and tested application aims to foster the comprehensive development of literacy through a playful and educational structure, based on five thematic worlds, each with a specific focus but all oriented towards the foundations of writing and reading (Hinojo-Lucena et al., 2023).

- **First world**: recognition of letters and basic literacy skills. In this world, students participate in activities such as solving word searches that contain vowels and consonants. These activities are designed to familiarise them with the alphabet and help develop fundamental reading and writing skills.
- **Second world:** image and word association. Here, students associate images of animals with the corresponding words. This activity promotes the relationship between images and written words, facilitating vocabulary learning and word comprehension in context.
- **Third world:** sound recognition. Students listen to audio clips and select the word corresponding to the heard sound. This approach helps develop the ability to recognise sounds and associate them with written words, thereby improving auditory and phonological comprehension.
- **Fourth world:** writing skills. This world focuses on writing. Students complete words with missing syllables and practice writing full words. These activities are designed to improve spelling, word formation, and writing accuracy.
- **Fifth world:** reading skills development. Centred on reading, this world offers students various reading activities tailored to their level. These activities aim to enhance fluency, comprehension, and reading speed, providing a solid foundation for the development of more advanced literacy skills.

A key aspect of AppRende is the use of gamification as an educational strategy. This is implemented through challenges that require successfully completing one level to advance to the next. For this reason, the worlds are initially locked with padlocks, motivating students to meet prior objectives to unlock them (Figure 3).

Additionally, upon completing each activity, users receive a score that reflects their performance, allowing both students and teachers to assess individual progress based on points earned from completing tasks and achieving goals.

With this interactive and game-based structure, the application was implemented as an educational intervention with the purpose of evaluating its effectiveness in developing and improving literacy skills.

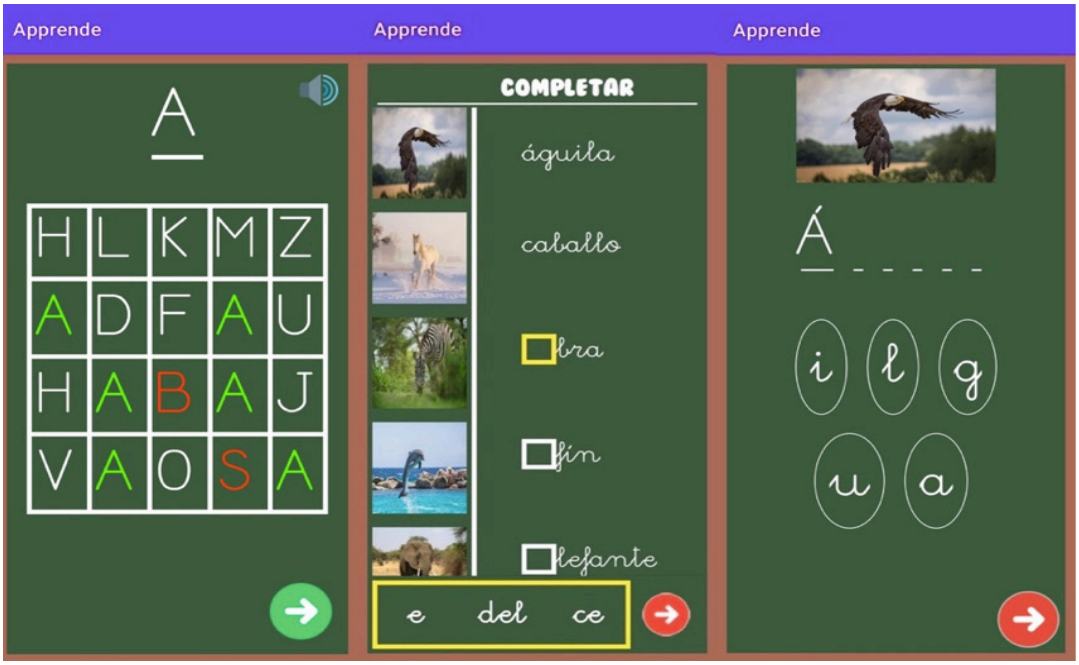
FIGURE 3. Apprende interface

Conversely, in the control group, participants were assigned to work on the same content as the experimental group, but using printed materials instead of the mobile application.

Regarding the teacher's role, two teachers managed the entire process and the intervention. The teachers' role during the intervention was to provide assistance when needed, handle technical issues with the tablet, and guide the content. Additionally, they reviewed the scores obtained since the application provided a score at the end of each activity and each world, except for the writing and reading activities in worlds two and three, where the teachers had to review and assess the correctness of the written words and the accuracy of reading each sentence.

The application is in a restricted access state, requiring communication with the main authors for authorisation to use it. Below are images of the application for illustration purposes (Figure 4).

FIGURE 4. Example of activities



3.4. Data analysis

In the data analysis, normality in the distribution of scores was calculated using the Kolmogorov-Smirnov test with Lilliefors significance correction. Subsequently, descriptive statistical values of means and standard deviations, Mann-Whitney U, p-value, and Cohen's d were calculated to test the hypothesis 1. For hypothesis 2, the decision was made to use the non-parametric Mann-Whitney U test (intergroup) and Wilcoxon signed-rank test (intragroup). The data analysis was conducted using SPSS statistical software v.25.

4. RESULTS

The Kolmogorov-Smirnov normality test with Lilliefors significance correction confirmed that the data did not follow a normal distribution, as the p-value was below .05 for both the pretest (K-S = .170; df = 43; p = .003) and the posttest (K-S = .232; df = 43; p = .000).

The results from the comparisons between student groups based on sex and age in the pretest phase (Table 1) and posttest phase (Table 2) showed no significant differences according to sex and age. Thus, hypothesis 1 was rejected. However, the highest means in the pretest were for the group of girls and seven-year-olds. In the posttest, the highest mean was for the group of boys and seven-year-olds.

TABLE 1. Differences between groups based on sex and age variables in the pretest measure

Data	n	M	DT	*u*	*p*	*d*
Sex						
Boy	28	642	1.83	216.0	.843	-.18
Girl	16	6.75	1.73			
Age						
6	28	6.28	1.86	176.0	.234	-.41
7	16	7	1.59			

TABLE 2. Differences between groups based on sex and age variables in the posttest measure

Socio-demographic data	n	M	DT	*u*	*p*	*d*
Sex						
Boy	28	8	1.12	208.0	.691	.31
Girl	16	7.5	1.93			
Age						
6	28	7.71	1.82	224.0	1	-.22
7	16	8	.36			

Regarding the intragroup comparisons (Table 3), the control group obtained a lower mean in the posttest compared to the pretest (.10), with no significant differences between the two measurement points (Z = -.872; p = .383). In contrast, the experimental group improved by more than two points compared to the pretest, showing statistically significant differences (Z = -4.327; p = .000).

TABLE 3. Intragroup comparisons on literacy competence

Group	n	M	DT	*z*	*p*	*d*
Control						
Pretest	20	6.80	1.96	-.872	.383	.05
Postest	20	6.70	1.32			
Experimental						
Pretest	24	6.33	1.63	-4.327	.000	-1.90
Postest	24	8.75	.76			

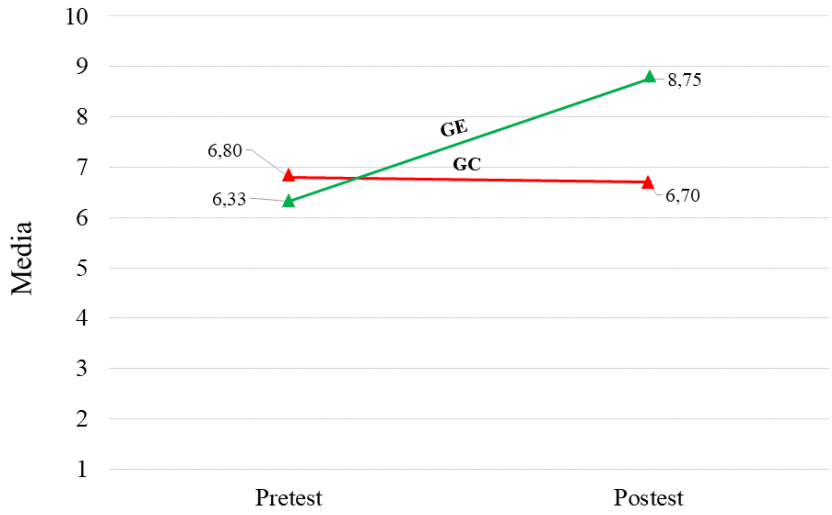
TABLE 4. Intergroup comparisons on literacy competence

Measure	n	M	DT	*u*	*p*	*d*
Pretest						
CG	20	6.80	1.96	184.0	.180	.26
EG	24	6.33	1.63			
Postest						
CG	20	6.70	1.32	16,0	.000	-1.90
EG	24	8.75	.76			

Furthermore, the comparison between the control group (CG) and the experimental group (EG) in the pretest and posttest measures (Table 4) indicated no significant differences in the pretest measure between the groups (p = .180). Thus, both groups started from a similar baseline, although the mean of the EG was lower. However, in the posttest measure, the EG obtained a higher mean than the CG, specifically by more than two points, showing significant differences between them (p = .000).

Finally, the average scores obtained in the posttest measure were higher than in the pretest measure for the EG compared to the CG (Figure 5). This establishes an improvement in literacy competence after completing the AppRende programme for migrant students.

FIGURE 5. Comparison of literacy competence levels between pretest and posttest measurements after the AppRende programme



5. DISCUSSION AND CONCLUSIONS

This study focused on evaluating the impact of a mobile application aimed at leveraging the advantages of technology to enhance the teaching and learning of literacy skills among migrant students. Mobile applications offer multiple advantages over traditional learning approaches, including the possibility to personalise learning, adapt to the individual needs of students, provide immediate feedback, and foster motivation and engagement. In this context, mobile applications have been shown to promote greater autonomy and adaptability, especially in multicultural settings where linguistic and cultural diversity play a critical role in students' integration (Criollo et al., 2021; Rosell-Aguilar, 2017). This combination of flexibility and adaptability makes mobile apps particularly effective tools in diverse and multicultural educational settings. Furthermore, the use of mobile devices is widely accepted and utilised by students, making mobile applications an accessible and familiar tool for their learning.

Consequently, the findings of this research reveal that the use of a mobile application, which addresses both reading and writing, had a positive impact on improving literacy skills in migrant students in Primary Education. It was observed that the experimental group (EG) showed a significant improvement in literacy competence after participating in the AppRende programme compared to the control group. In the pretest measure, the experimental group achieved an average of 6.33, while in the posttest measure, the average increased to 8.75. This difference of more than two points was statistically significant ($p < .001$) and resulted in a substantial improvement in the literacy competence of the migrant students. This is consistent

with findings from previous studies on mobile literacy interventions, such as those of Jiménez-García and Martínez-Ortega (2017), who also observed significant improvements in literacy skills through mobile-based instruction.

This evidence highlights the importance of incorporating mobile technology to address the specific literacy needs of migrant students, who often face challenges such as language barriers and cultural adaptation (Abad-Quintanal, 2022).

On the other hand, when comparing the control group (CG) and the experimental group in the posttest measure, it was found that the experimental group obtained a higher average than the control group. The average of the control group in the posttest measure was 6.70, while that of the experimental group was 8.75, showing a significant difference between both groups ($p < .001$). This result confirms that mobile learning tools are not only effective but offer distinct advantages over traditional methods of literacy instruction, as shown by earlier studies (Calvo, 2019; Fernández-Batanero et al., 2022). Similarly, Forné et al. (2022) concur in their findings, observing notable progress in the reading skills of the group of children who participated in the phonological training programme. This comparison aligns with our results, reinforcing the value of targeted literacy interventions.

The statistical methods used—specifically, the pre-test-post-test control group design—provided a reliable framework for assessing the impact of the intervention. This allowed for clear comparisons between groups and reinforced the validity of the study's conclusions. The significant differences found between the experimental and control groups align with similar research supporting the effectiveness of mobile applications in literacy education (Máñez-Carvajal & Cervera-Mérida, 2022).

It is important to consider the role of cultural and linguistic factors in literacy acquisition, as noted by authors like Bai (2019) and González-Monteagudo and León-Sánchez (2020). Proficiency in the host country's language plays a critical role in the social inclusion of migrant students, and tools like the AppRende programme help bridge this gap by offering accessible literacy support tailored to their needs. While it has been observed that migrant students often exhibit lower academic performance (Abad-Quintanal, 2022), this study demonstrates that mobile applications can be effective interventions for reversing this trend. This finding is in line with previous studies, such as those by Tajic and Bunar (2023), which emphasize the necessity of addressing literacy as a core component of migrant student integration.

Moreover, these students experience higher levels of anxiety and confusion in the school environment (El Jouhri et al., 2022), and the introduction of technology-based literacy interventions has also been shown to reduce anxiety by providing a more engaging and student-centred learning experience. By reducing anxiety and providing more tailored literacy support, mobile learning tools can help lower the dropout rates among migrant students, which remains a critical issue, as highlighted by Ceballos-Vacas and Trujillo-González (2021).

The interpretation of these findings highlights the importance of effectively addressing the literacy of the migrant population in Primary Education. It is essential to emphasise the need for further research to strengthen the link between literacy interventions and their long-term effects on academic performance and social integration. Educational policies must increasingly focus on creating inclusive environments

where mobile technologies are integrated into curricula, providing a seamless and supportive framework for migrant learners (UNESCO, 2022).

Teacher training in interculturality and the pedagogical use of digital tools is also a crucial factor in maximizing the benefits of mobile literacy applications (Biasutti et al., 2021). These actions will contribute to their academic success and integration into the educational environment. By fostering a more inclusive approach to education, educators can ensure that migrant students are better prepared to thrive both academically and socially.

Ultimately, comprehensive literacy development is not merely an academic achievement but a means of fostering social inclusion and effective participation in broader societal contexts (Acedo-Tapia & Maqueda-Berrocal, 2022). In the context of migrant students, literacy competence becomes even more relevant. These students often face additional challenges, such as learning a new language and adapting to a new culture. The acquisition of the host language is essential for their academic success and integration into society. Thus, literacy education serves as a bridge for migrant students, enabling them to fully engage with both the academic content and their peers.

Thus, the study's objective was achieved by providing empirical evidence on the effectiveness of this technological tool, determining whether it had a positive impact on improving the literacy competence of migrant students compared to a control group, supporting the proposed hypothesis. The results indicate that mobile applications are not only effective tools for literacy development but also offer scalable and adaptable solutions that can be implemented across different educational settings to meet the needs of diverse student populations.

5.1. Limitations and future lines of research

Regarding the limitations, the sample used in the study was small, which may restrict the generalisation of the results to a broader population. The lack of diversity in terms of ethnic origin and cultural backgrounds can also influence the applicability of the findings. On the other hand, although a quasi-experimental design with a control and experimental group was used, this type of design does not allow for definitive causal relationships to be established. There may be other uncontrolled factors that could influence the observed results.

Finally, as practical implications and future lines of research: (i) the results support the effectiveness of using mobile applications specifically designed to support the literacy skills of migrant students. This implies that educational app developers can focus on creating technological tools that adapt to the linguistic and educational needs of this population; (ii) the use of mobile applications can provide additional and specific support to migrant students in their literacy learning. These applications can adapt to different levels of linguistic competence and offer interactive and personalised activities to help students improve their skills effectively; (iii) conducting comparative studies that compare the effectiveness of the mobile application with other interventions designed to support the linguistic development of migrant students. This could include traditional paper-based interventions, individualised tutoring programmes, or other technological tools; (iv) although the study shows that the designed mobile application was effective, additional research

can be conducted to improve and optimise the application. This could include incorporating additional features, adapting to different languages and levels of linguistic competence, and personalisation to meet the individual needs of migrant students; (v) to better understand the long-term impact of the mobile application on the linguistic development of migrant students, research can be developed that conducts long-term follow-ups of the participants. This would allow for the evaluation of whether the observed improvements are maintained over time and if they have a lasting effect on the students' literacy skills.

In summary, this work provides a solid foundation for implementing mobile applications and appropriate pedagogical strategies that support the linguistic learning and development of migrant students in the context of Primary Education. These implications can guide educators, app developers, and educational policymakers in creating and promoting inclusive and effective learning environments for migrant students.

6. FUNDING

This work has been financed with public funds through the project 'AppRende: Inclusion and Improvement of the Reading and Writing Competence of Immigrant Students in Andalusia' (Reference: PY20_00077), from the call for grants for I+D+i projects on a competitive basis within the scope of the Andalusian Plan for Research, Development and Innovation (PAIDI 2020).

7. REFERENCES

Abad-Quintanal, G. (2022). El desconocimiento del español por los inmigrantes y su integración en el ámbito educativo: desafíos, conflictos potenciales y respuestas. *Lengua y migración*, *14*(1), 89-109. https://doi.org/10.37536/LYM.14.1.2022.1597

Acedo-Tapia, E. M., & Maqueda- Berrocal, R. (2022). La poesía como método para trabajar la lectoescritura en educación primaria: Proyecto educativo a través de Gloria Fuertes. *Visual Review. International Visual Culture Review / Revista Internacional De Cultura Visual*, *11*(3), 1–10. https://doi.org/10.37467/revvisual.v9.3668

Bai, H. (2019). Preparing teacher education students to integrate mobile learning into elementary education. *TechTrends*, *63*(6), 723-733. https://doi.org/10.1007/s11528-019-00424-z

Bayón-Calvo, S. (2019). Una radiografía del abandono escolar temprano en España: Algunas claves para la política educativa en los inicios del siglo XXI. *Revista complutense de educación*, *30*(1), 35-53. http://dx.doi.org/10.5209/RCED.55465

Biasutti, M., Concina, E., Frate, S., & Delen, I. (2021). Teacher professional development: Experiences in an international project on intercultural education. *Sustainability, 13*(8), 1-16. https://doi.org/10.3390/su13084171

Cáceres-Reche, M. P., Romero-Rodríguez, J. M., de la Cruz-Campos, J. C., & Berral-Ortiz, B. (2024). Mejora de la competencia lectoescritora del alumnado migrante por medio de herramientas digitales. *Campus Virtuales, 13*(1), 83-91. https://doi.org/10.54988/cv.2024.1.1316

Calvo, V. (2019). El diario de lectura en los procesos de acogida de familias inmigrantes. *Ocnos. Revista De Estudios Sobre Lectura, 18*(1), 41-51. https://doi.org/10.18239/ocnos_2019.18.1.1789

Ceballos-Vacas, E., & Trujillo-González, E. (2021). Dificultades y apoyos emocionales del alumnado migrante: un estudio de caso en un instituto multicultural. *Aula Abierta, 50*(4), 767–776. https://doi.org/10.17811/rifie.50.4.2021.767-776

Chou, P. N., & Feng, S. T. (2019). Using a Tablet Computer Application to Advance High School Students' Laboratory Learning Experiences: A Focus on Electrical Engineering Education. *Sustainability, 11*(2), 381. https://doi.org/10.3390/su11020381

Criollo, S., Guerrero-Arias, A., Jaramillo-Alcázar, Á., & Luján-Mora, S. (2021). Mobile learning technologies for education: Benefits and pending issues. *Applied Sciences*, *11*(9), 4111. https://doi.org/10.3390/app11094111

EDUCASE (2022). *Horizon Report. Teaching and Learning Edition.* EDUCASE. https://library.educause.edu/resources/2022/4/2022-educause-horizon-report-teaching-and-learning-edition

El Jouhri, N., Bagur, S., Verger, S., & Rosselló-Ramon, MR (2022). El alumnado inmigrante y la mediación lingüística infantojuvenil. *Magis. Revista Internacional de Investigación en Educación*, *15*, 1-25. https://doi.org/10.11144/Javeriana.m15.aiml

Escarbajal, A., Navarro, J., & Arnaiz, P. (2019). El rendimiento académico del alumnado autóctono y de origen inmigrante en la Región de Murcia. *Tendencias Pedagógicas, 33*, 5–17. https://doi.org/10.15366/tp2019.33.001

Fernández-Batanero, J.M., Montenegro-Rueda, M., Fernández- Cerero, J.F., & Román-Gravan, P. (2022). Impacto de las TIC en las habilidades de escritura y lectura: una revisión sistemática (2010-2020). *Texto Livre*, *14*, 1-12. https://doi.org/10.35699/1983-3652.2021.34055

Ferreiro, E. & Teberosky, A. (1979). *Los sistemas de escritura en el desarrollo del niño*. Siglo XXI Editores.

Forné, S., López-Sala, A., Mateu-Estivill, R., Adan, A., Caldú, X., Rifà-Ros, X., & Serra-Grabulosa, J. M. (2022). Improving Reading Skills Using a Computerized Phonological Training Program in Early Readers with Reading Difficulties. *International Journal of Environmental Research and Public Health, 19*(18), 11526. https://doi.org/10.3390/ijerph191811526

García-Rodríguez, A., & Gómez-Díaz, R. (2016). Niños y apps: aprendiendo a leer y escribir en digital. *Alabe Revista De Investigación Sobre Lectura Y Escritura,* (13), 1-23. https://doi.org/10.15645/Alabe2016.13.6

González-Monteagudo, J., & León-Sánchez, M. (2020). Hacia una educación intercultural inclusiva. El caso del alumnado inmigrante en la escuela en España. En M. A. González-González (Ed.), *Migraciones, rasgaduras humanas* (pp. 131-167). Universitat Rovira i Virgili, Universidad de Manizales y Horizontes Humanos.

Hautala, J., Heikkila, R., Nieminen, L., Rantanen, V., Juha-Matti, L., & Richardson, U. (2020). Identificación de dificultades de lectura mediante una tecnología de evaluación basada en juegos digitales. *Revista de Investigación en Computación Educativa, 58*(5), 1003-1028. https://doi.org/10.1177/073563120905309

Hernández-Campos, E., Jaimez-González, C. R., & García-Mendoza, B. (2020). Interactive Mobile Applications to Support the Teaching of Reading and Writing of Spanish for Children in Primary Education. *International Journal of Interactive Mobile Technologies (iJIM), 14*(14), 64–79. https://doi.org/10.3991/ijim.v14i14.14925

Hinojo-Lucena, F. J., Aznar-Díaz, I., & Romero-Rodríguez, J. M. (2020). Mobile learning en las diferentes etapas educativas. Una revisión bibliométrica de la producción científica en Scopus (2007-2017). *Revista Fuentes, 22*(1), 37-52. https://doi.org/10.12795/revistafuentes.2020.v22.i1.04

Hinojo-Lucena, F. J., Fernández-Martín, F. D., Cáceres-Reche, M. P., & Berral-Ortiz, B. (2023). Innovación educativa en acción: La aplicación Apprende para el mejoramiento de la lectoescritura en contextos multiculturales. En S. Alonso-García, J. M. Romero-Rodríguez, M. N. Campos-Soto, & L. Hinojo-Cirre (Eds.), *Investigación educativa como factor de avance en las aulas 5.0* (pp. 11-24). Editorial Dykinson.

Jiménez-García, M., & Martínez-Ortega, M. Á. (2017). El uso de una aplicación móvil en la enseñanza de la lectura. *Información tecnológica, 28*(1), 151-160. http://dx.doi.org/10.4067/S0718-07642017000100015

Junta de Andalucía (2011). *Prueba de evaluación inicial de diagnóstico competencia lingüística.* https://shorturl.at/aqVX1

Máñez-Carvajal, C., & Cervera-Mérida, J. F. (2022). Desarrollo de aplicación móvil para niños con dificultades de aprendizaje de la lectura y escritura. *Información tecnológica*, *33*(1), 271-278. http://dx.doi.org/10.4067/S0718-07642022000100271

Martínez-Acosta, D., Suárez-Brieva, E., & Gordon-Hernández, Y. (2022). Aplicación móvil como estrategia de enseñanza para

iniciar el proceso de lectura a estudiantes en condición de discapacidad auditiva. *Información tecnológica, 33*(4), 1-12. http://dx.doi.org/10.4067/S0718-07642022000400001

Rodríguez, S., Valle, A., Gironelli, L. M., Guerrero, E., Regueiro, B., & Estévez, I. (2020). Performance and well-being of native and immigrant students. Comparative analysis based on PISA 2018. *Journal of adolescence, 85*, 96–105. https://doi.org/10.1016/j.adolescence.2020.10.001

Rosell-Aguilar, F. (2017). State of the App: A Taxonomy and Framework for Evaluating Language Learning Mobile Applications. *CALICO Journal, 34*(2), 243–258. https://www.jstor.org/stable/90014690

Tajic, D., & Bunar, N. (2023). Do both 'get it right'? Inclusion of newly arrived migrant students in Swedish primary schools. *International Journal of Inclusive Education*, *27*(3), 288-302. https://doi.org/10.1080/13603116.2020.1841838

Toledo-Vega, G., Quilodrán, F., & Cerda, K. (2022). Alfabetización escolar en aprendientes de lenguas adicionales no alfabetizados: aspectos sociales, cognitivos, lingüísticos y pedagógicos. *Lengua y migración*, *14*(2), 129-149. https://doi.org/10.37536/LYM.14.2.2022.1386

UNESCO. (2022, September 29). *Mejores prácticas de aprendizaje móvil*. UNESCO. https://www.unesco.org/es/digital-education/mobile-learning-practices

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. Junio 2025 - pp. 108-133 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20521

The effectiveness of Grammarly application and teacher feedback for undergraduate EFL students' writing skills

La eficacia de la aplicación de Grammarly y la retroalimentación proporcionada por el docente en relación con los estudiantes de EFL en el ámbito de la escritura

RECEIVED 25/11/2024 ACCEPTED 24/02/2025 PUBLISHED 01/06/2025

Delsa Miranty
ORCID 0000-0003-3466-9127
Universitas Sultan Ageng Tirtayasa, Indonesia
delsa12@untirta.ac.id

Utami Widiati
ORCID 0000-0002-8603-4556
Universitas Negeri Malang, Indonesia
utami.widiati.fs@um.ac.id

Bambang Yudi Cahyono
ORCID 0000-0001-5210-5208
Universitas Negeri Malang, Indonesia
Bambang.yudi.fs@um.ac.id

Tengku Intan Suzila Tengku Sharif
ORCID 0000-0002-2529-9112
Universiti Teknologi MARA, Malaysia
Intansuzila@uitm.edu.my

ABSTRACT

This research investigates the efficacy of Grammarly as an application and teacher feedback in enhancing the writing skills of Indonesian undergraduate EFL students. The Switching Replication Design encompasses 78 first-year students enrolled in a writing course. The study examines the impact of these feedback methods on writing, skill enhancement, and student responses. Data is gathered via pre-and post-tests and a closed-ended questionnaire and analyzed using One-Way ANCOVA and descriptive statistics in SPSS 20. Research shows that Grammarly, when combined with teacher feedback, enhances diction, language usage, and mechanics and from teacher feedback has more effect on content and organization. The initial experimental group surpasses the second, attaining a moderate N-Gain score. Students react favorably to both feedback modalities, with Grammarly commended for its automated recommendations. The research concludes that integrating Grammarly and teacher feedback significantly improves writing skills, particularly in diction, language usage, mechanics, content, and organization.

KEYWORDS Applications; Grammarly; N-Gain score; One-Way ANCOVA; teacher feedback.

RESUMEN

Este estudio examina la eficacia de Grammarly como una aplicación y una retroalimentación pedagógica en la mejora de la escritura de los estudiantes de EFL en Indonesia. La disciplina de la Switching Replication Design abarca 78 estudiantes de primer año que están inscritos en un curso de escritura. La investigación analiza la repercusión de estos métodos de retroalimentación en la escritura, el desarrollo de habilidades y las respuestas de los estudiantes. La información se recaba a través de pre- y post-tests, así como a través de un questionnaire closed-ended, y se analiza mediante el uso de One-Way ANCOVA y descriptive statistics en SPSS 20. Se ha demostrado que el uso de Grammarly, cuando se combina con la retroalimentación del profesor, potencia la diction, el uso del lenguaje y los mecanismos, y la retroalimentación del profesor tiene una mayor influencia en el contenido y la organización. El grupo experimental inicial supera al grupo experimental secundario, logrando una puntuación moderada de N-Gain. Los estudiantes responden favorablemente a ambas modalidades de retroalimentación, con Grammarly siendo elogiado por sus consejos automatizados. La investigación concluye que la integración de Grammarly y la retroalimentación pedagógica contribuye significativamente a la mejora de las habilidades de escritura, especialmente en aspectos como la diction, el uso del lenguaje, las técnicas, el contenido y la organización.

PALABRAS CLAVE Aplicaciones; Grammarly; puntuación N-Gain; ANCOVA unidireccional; retroalimentación del docente.

1. INTRODUCTION

Writing in English is essential for foreign language learners. To produce a piece of writing, they should use appropriate vocabulary, form meaningful sentences of different text types, and choose proper arrangements to link sentences (Nasser, 2018) through the writing processes of planning and organizing ideas followed by drafting, reviewing, and editing to improve the writing quality before arriving at the final version (Albesher, 2022; Bai et al., 2022; Martínez et al., 2020; Mirzae & Shamsudin, 2023; Teng et al., 2022). Technology must be utilized in writing classes to deal with the writing processes (Aljameel, 2022; Loncar et al., 2023; Taskiran & Goksel, 2022; Wong et al., 2022; Wu, 2022; Yuniar et al., 2019; Zhang & Zou, 2022).

Recent technological advances in artificial intelligence (AI) have paved the way for improved electronic writing tools, as well as the development of innovative ones. These writing support systems provide human-like sentence completion and text generation, making them essential for many writers and students (Alharbi, 2023). Alharbi showed that students use more AI-powered writing tools to improve their writing. Other studies found that deep learning may improve human-AI learning community performance (Song & Song, 2023; Wang et al., 2023). In sum, new AI-powered writing tools have improved writing performance.

Technology has been a learning tool for electronic corrective feedback by EFL teachers to their students' compositions (Almusharraf & Alotaibi, 2023; Barrot, 2020; Lei, 2020; Liu et al., 2023; Miranty et al., 2023; Mohsen & Alshahrani, 2019; Parra & Calero, 2019; Tang & Rich, 2017; Zou et al., 2023), fostering EFL students' writing achievement (Andina et al., 2020). One of the new language-learning technologies is Automated Writing Evaluation (AWE), an outstanding support for meeting diagnostic feedback on writing aspects (Ranalli et al., 2017). Bai and Hu (2017) stated that AWE offers users consistent explanations and immediate feedback. These two characteristics have allowed students to overcome time constraints, develop writing skills at their own pace, and participate and interact independently in language classes (Liao, 2016; Zhang et al., 2020). In addition, the consistency and objectivity of AWE allow students to improve their writing mechanics and accuracy (Tian & Zhou, 2020).

In writing courses, students can use AWE tools to plan, write, get automated feedback, revise, and improve their writing (Roscoe et al., 2017). AWE can also promote learner autonomy (Tang & Rich, 2017) and save teachers' time (Palermo & Wilson, 2020). Moreover, students can access sample writing and online dictionaries through some AWE tools, and teachers may have access to additional tools like plagiarism detection (Alharbi, 2023; Ariyanto et al., 2021; Barrot, 2023; Hockly, 2019; Joo, 2021; Shadiev & Feng, 2023).

Much literature has shown that the use of Grammarly, one of the AWE tools, helps students understand grammatical rules and their writing tasks (Agustin & Wulandari, 2022; Cavaleri & Dianati, 2016; Estacio et al., 2022; Fitriana & Nurazni, 2021). Grammarly can correct grammaticality, assess the correctness and readability of writings, and suggest vocabulary enhancements (Lei, 2020). It also has a built-in plagiarism detector and can check styles beyond sentences (Grammarly, 2019). Thus, the core concept to use Grammarly is that electronic corrective feedback can help EFL students improve their writing.

The use of Grammarly has been empirically proven beneficial. Grammarly encourages learner autonomy by requiring students to evaluate their work using the system's feedback (Qassemzadeh & Soleimani, 2016). Cavaleri and Dianati (2016) and O'Neil and Russel (2019) found that students of both studies liked the easy-to-use and valuable feedback of Grammarly. It can deal with writing complexity in different genres, allowing teachers to spend more time with students and providing direct, indirect, and metalinguistic feedback (Bailey & Lee, 2020). O'Neill and Russell (2019) found Grammarly valuable because it allows students to choose their preferred feedback strategy based on their needs. Using Grammarly software in EFL writing helps students make fewer errors in terms of vocabulary usage, language usage, and mechanics (Ghufron & Rosyida, 2018). Grammarly can be a proofreading service for spelling, punctuation, vocabulary, and plagiarism (Barrot, 2023; O'Neil & Russel, 2019). In addition, Grammarly has proved helpful for L2 writers right after they finish their first draft because the feedback lets them see where they might need to improve their writing before turning it into the teacher (Almusharraf & Alotaibi, 2023). In sum, Grammarly is a useful tool; it provides students with some autonomy and motivation elements necessary in writing classes.

However, studies have found different results and essential gaps in our understanding of how Grammarly can help improve writing skills. Grammarly adds to teacher feedback without clarifying who is responsible for fixing higher- and lower-order issues (Koltovskaia, 2023). Therefore, it is still unclear how well this integration improves writing quality overall. Research comparing Grammarly's corrections with teacher feedback shows how valuable and vital the tool is, suggesting that Grammarly cannot replace human evaluation (Khushk et al., 2024). Grammarly considerably reduces mistakes, but how well it works depends on how well the learners already know the language (Jomaa & Jibroo, 2024). These results show that more research is needed to fully understand Grammarly's role in various settings and its relations with teacher feedback.

In this study, Grammarly is complemented with teacher feedback, which educators have used in workbooks, exams, and lessons. Teacher corrective feedback impacts content and organization more than diction, language use, and mechanics (Ghufron & Rosyida, 2018). The best-known teacher feedback is written feedback on students' writing pieces (Sermsook et al., 2017). In sum, teacher feedback has been viewed as an essential part of improving the performance of L2 writers. The present study explores how Grammarly's automated feedback can be combined with teacher feedback. In Indonesia's collectivist culture, where teacher-student trust and respect affect learning outcomes, the study stressed the relational aspect

of feedback. Boud and Dawson (2023) found that teacher feedback literacy—the ability of teachers to design and deliver feedback—can improve learning, as also highlighted by Heron et al. (2023).

Zheng et al. (2023) showed that verification, scaffolding, and teacher praise improved SRL strategies. EFL education in Indonesia emphasizes independent learning, making this finding relevant. In Indonesia, feedback delivery may be inconsistent, especially in large classrooms. Also, teacher feedback should balance praise and constructive criticism for writing development. In conclusion, this feedback from Grammarly and teacher feedback may improve writing development.

Studying how students interact with teachers and automated feedback tools like Grammarly has expanded research on feedback's impact on writing. One study found that Hungarian university students engaged moderately to low with teachers and automated feedback, focusing on form rather than meaning, which affected their writing revision strategies (Thi et al., 2022). Then, research on AWE tools like Grammarly, ChatGPT, and Quillbot shows that they improve students' feedback literacy, depending on their feedback-seeking behaviors (Gozali et al., 2024). Moreover, using these technologies in writing courses has necessitated frameworks to help educators integrate AWE tools while considering digital literacy and ethical AI use (Arredondo & Laurens, 2023). New technologies like the pandemic-era metaverse could improve educational delivery by creating virtual learning ecosystems that change how students interact with feedback and digital tools.

This study employed a Switching Replication Design in two writing classes, providing two different types of feedback, Grammarly and teacher feedback, in a different order, a research design different from the previous studies about Grammarly and teacher feedback, such as the ones by Karyuatry et al. (2018), Ventayen and Ventayen (2018) and ONeill and Russell (2019).

In the context of teaching argumentative writing, many studies examined the effect of various strategies and methods on argumentative writing skills (Awada & Diab, 2023; Backman et al., 2023; Jumariati & Sulistyo, 2017; Landrieu et al., 2023; Olson et al., 2023; Wang & Chiu, 2024). They discovered that EFL student teachers with limited writing experience struggled with vocabulary and grammar when composing an argumentative writing task (Ghufron & Rosyida, 2018). The most consistent finding in these studies is that students, regardless of their stage of development, struggle to write argumentative texts and that it is necessary to provide planned and educational tools to ensure student's success in the writing course. Using AWE may help overcome students' problems in writing (Fitria, 2021). Grammarly students report increased confidence and writing quality due to the grammar and spelling correction features (Setyani et al., 2023). Despite some inaccuracies in certain features, such as plagiarism detection, students find the app helpful in increasing their self-awareness and vocabulary (Setyowati et al., 2024). Less-proficient EFL students mainly show positive engagement with Grammarly despite their low cognitive engagement (Anastasia et al., 2024). This suggests that integrating Grammarly into the writing process, in addition to traditional teacher feedback, significantly impacts students' writing skills (Inayah & Apoko, 2024). In sum, Grammarly's effectiveness in improving writing skills is evident, especially with proper guidance and an understanding of its limitations.

However, only a few studies have been done to know the effectiveness of AWE for undergraduate students, Grammarly in particular as electronic feedback, and teacher feedback on student writing performance for Indonesian undergraduate EFL students (Fahmi & Cahyono, 2021; Karyuatry et al., 2018; Miranty

& Widiati, 2021). Therefore, this study aims to determine the effectiveness of Grammarly followed by teacher feedback in writing classes of Indonesian undergraduate EFL students with two experimental courses. Three research questions are formulated in this study:

1. How does Grammarly followed with teacher feedback affect the writing skill of Indonesian undergraduate EFL students?
2. To what extent does Grammarly followed with teacher feedback affect the writing skill of Indonesian undergraduate EFL students?
3. How do the Indonesian undergraduate EFL students respond to the use of Grammarly followed with teacher feedback?

2. LITERATURE REVIEW

2.1. Automated writing evaluation (AWE) as technology in the class

With advancements in educational technology and a growing reliance on technology, several studies demonstrated the efficacy of using AWE in the L2 classroom. AWE enables different technology interactions, learners, teachers, and peers to write. For instance, students can use the system to plan, write, receive automated feedback, review their work, and improve their writing (Wilson & Roscoe, 2019). AWE significantly improved students' writing quality and genre, with argumentative writing outperforming academic and mixed writing genres (Zhai & Ma, 2023).

AWE is adapting to various cultural practices as a well-known technology and minimizing teacher writing. AWE enables students to adjust their writing style to specific cultural practices, leading to improved writing skills (Zhai & Ma, 2023). Stevenson (2016) analyzed how AWE is used in the writing classroom as a teaching method and summarizes AWE's use in class. AWE is a popular educational technology that allows teachers to write less (Roscoe et al., 2017). Next, AWE is more effective than traditional methods for developing writing skills, including classroom, teacher-led, and peer interaction (Li, 2023). Moreover, in academic writing, one effective strategy for encouraging students to take responsibility for their improvement could be to emphasize the constant need to draft and revise. In sum, technology is needed as feedback for the students' writing texts, and many teachers should use electronic feedback, such as Grammarly.

2.2. Sources of feedback

Generally, feedback encompasses corrective feedback, focusing on formal aspects of learners' language and improving linguistic accuracy. Computer-generated information was used to deliver it automatically. In this study, the sources of feedback are electronic feedback (Alsmari, 2019; Barrot, 2020; Mohsen & Alshahrani, 2019) and teacher feedback (Lie & He, 2017; Wu et al., 2023; Yang et al., 2023).

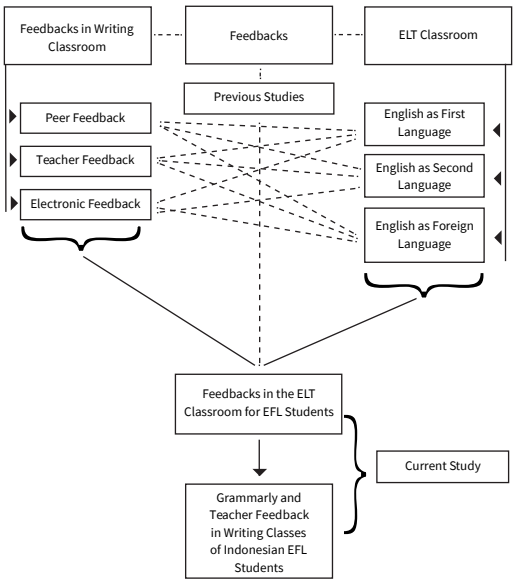
Electronic feedback is a strategy to provide automated feedback from the computer to draw attention to students' writing. Furthermore, one electronic feedback method that can be applied to students' texts is Grammarly. Grammarly is a popular educational technology that eliminates the need for teachers to write, and it helps students enhance their writing (Fitriana & Nurazni, 2022; Roscoe et al., 2017). This is in line with

the study conducted by several researchers (Dewi, 2023; Karyuatry et al., 2018; O'Neil & Russel, 2019), who also found that students perceived Grammarly as a powerful tool for quickly checking for possible stylistic errors in grammar. Additionally, they discovered that students were more satisfied with Grammarly's feedback than with teacher feedback. As a result, the writing quality can be improved. Then, numerous studies have confirmed Grammarly's beneficial effect (Ebyary & Windeatt, 2010; Fitria, 2022; Liao, 2016; Saman et al., 2023). In sum, Grammarly provides electronic writing feedback to students and is helpful for quickly checking grammar and stylistic errors.

Teacher feedback is commonly viewed as an essential part of improving the performance of L2 writers. The best-known teacher's feedback is written feedback on students' writing pieces (Sermsook et al., 2017). Teachers must do a needs assessment and method analysis on the components of writing, the amount, and how feedback should be provided to provide a source of feedback. Many of the studies on feedback focus on teacher feedback, with the study's primary object of giving feedback and the particular type of feedback's efficacy (Elumalai, 2019; Pearson, 2018). Moreover, the written feedback from the teacher needs to be easy to follow, practical, precise, and easy to understand (Kurnia, 2022). Teachers' written feedback is concise, focused, practical, and crucial for enhancing L2 writing proficiency.

Grammarly and teacher feedback work well together to address various aspects of writing, particularly lower-order concerns (LOCs) and higher-order concerns (HOCs). Grammarly provides immediate and consistent corrections for mechanical issues like grammar, spelling, and vocabulary, helping students improve their writing skills (Ayan & ErdemİR, 2023). However, it struggles to provide meaningful feedback on complex areas such as content, organization, and coherence, which require nuanced and context-specific insights (Bulatović et al., 2024). In contrast, teacher feedback is critical in guiding students through higher-order aspects, providing detailed advice on argument structure, idea development, and logical flow (Shum et al., 2023). While Grammarly promotes mechanical accuracy, teacher feedback is critical for improving the depth and clarity of content and organization. In conclusion, while Grammarly effectively handles technical corrections, teacher feedback remains critical for fine-tuning content and structure.

FIGURE 1. Theoretical Framework of Grammarly in ELT Classroom



As shown in Figure 1, the reason for using Grammarly in this study is that Grammarly focuses on academic writing in higher education. This tool combines machine-based artificial intelligence and natural language processing with deep learning algorithms to deliver real-time, quick, comprehensive writing evaluation results. Teacher feedback is crucially needed to fulfil the interactions of human beings among the students, which cannot be replaced with machines.

2.3. Argumentative text writing

The essay is a standard writing unit in college writing classes, with the argumentative essay being the most common type. The writer defends their position and attempts to rationalize the counter-arguments in the argumentative texts (Balta, 2018). Then, an argumentative essay is a piece of writing that argues for one's position while refuting opposing viewpoints (Özdemir, 2018). Next, in argumentative writing, writers can convey their opinions but should use objective sources (Landrieu et al., 2022). As a result, argumentative essays rely on cohesion to connect sentences, which assist students in developing coherent arguments for scientific papers.

The most crucial general skill in higher education is writing an argumentative essay. Students in higher education need even more guidance regarding their academic writing, primarily argumentative writing (Kleemola et al., 2022). Then, first-year students in higher education must learn argumentative writing to meet their educational requirements (Ghanbari & Salari, 2022). Moreover, idea generation, topic-oriented writing, and learner autonomy are essential components of argumentative writing for university students (Wu & Wang, 2023). In conclusion, argumentative writing is a critical competence in university students, and in addition to meet academic requirements, college freshmen must learn how to write argumentative writing along with get feedback concerning their argumentative writing.

In this study, the students were asked to write argumentative texts with topics given by the teacher. Through argumentative writing, the students were asked to provide their opinions, criticism, and ideas related to the issue and supported by evidence and facts. Moreover, the students were allowed to integrate using Grammarly, followed by teacher feedback.

3. METHOD

3.1. Research Design

This study employed a Switching-Replications (SR) design, a robust hybrid experimental framework, to compare Grammarly's electronic and teacher feedback on undergraduate EFL students' writing skills (Williams & Lowman, 2018). William and Lowman add that SR is one of the most potent hybrid experimental designs. The design split English Education Department students into two experimental groups (G1 and G2), with both groups receiving SR. The SR design ensured all participants received feedback that prevented the first experimental group from becoming hostile, reducing social threats to internal validity. There was no control group, so the first and second experimental groups served as the main subjects. The first experimental group (G1) received both Grammarly and teacher feedback, while G2 received teacher and Grammarly feedback, as shown in Table 1.

TABLE 1. Switching replication design

Group	Pre-test	Independent Variable 1	Independent Variable 2	Post-test
G1	Y1	X	O	Y2
G2	Y1	O	X	Y2

G1: First Experimental Group (G1)
G2: Second Experimental Group (G2)
Y1: Pre Test
Y2: Post Test
X: Grammarly
O: Teacher Feedback

As shown in Table 1, the Switching Replication Design was applied in this study with feedback from Grammarly and teacher feedback for two experimental groups, but with a different order of feedback when given to the students. The first experimental group started with Grammarly, followed by teacher feedback, but the second one started with teacher feedback, followed by Grammarly. In sum, both types of feedback were given to the students but in a different order to determine their effectiveness in improving the undergraduate EFL students' writing skills.

3.2. Population and sample

This research was done at a public university in Banten, Indonesia, where English is not spoken. First-year writing students from the 2020/2021 Teachers' Training and Education Faculty were studied. We sampled intact classes from two classes of 78 students: 40 in the experimental class and 38 in the other. Research design and statistics determined the sample size.

This study recruited Grammarly-using EFL writing students who were willing to participate. Next, this study included 78 students from two classes: 40 in the first experimental group and 38 in the second. Cluster sampling selected 78 second-semester students from 113. This study required (1) Grammarly-taught EFL writing class enrollment and (2) willingness to participate. The reason for involving the students was that one of us (the first author) was one of the Department's faculty members. Therefore, it was easy to get access to the students for conducting the study and distributing the closed-ended online questionnaire.

3.3. Research instruments

In this study, validators verified the data from tests and questionnaires. Two writing tests were given before and after treatments in the study. Argumentative writing tests would demonstrate students' writing skills in this study. Students wrote two argumentative texts on the topics for this test (Yamanishi et al., 2019). This rubric focused on the content, organization, diction (choice of words), language use, and mechanics. Then, it was modified in this study. These profiles were created in 1981 by Jacobs and his colleagues (Jacobs et al., 1981). Essays are graded on a 100-point scale for content (30), organization (20), vocabulary (20), language use (25), and mechanics (30) (5 points). Each set of criteria generates interval scores.

The study used a closed-ended online questionnaire with a 5-point Likert scale (strongly disagree to agree strongly) to address the third research question on students' responses to Grammarly. The questionnaire, consisting of 25 questions on effectiveness, source-based writing instruction, and usage frequency, was distributed via Google Forms in both English and Bahasa Indonesia to ensure clarity. After a validator review and suggestions, the questionnaire was linked to Grammarly's effectiveness in Indonesian undergraduate EFL writing classes. The reliability and validity were assessed through Pearson correlation, Cronbach's alpha, and SPSS 20, which also calculated test normality and homogeneity. Moreover, the closed-ended questionnaire was adapted from (Parra & Calero, 2019; Zhang et al., 2020).

3.4. Data collection procedure

This study involved first-year English Education students at a public university in Banten, Indonesia, from April to August 2021. It included a pre-test, four assignments on different topics, a post-test, and a questionnaire. The two experimental groups received Grammarly and teacher feedback but in different orders. The study aimed to assess whether Grammarly, followed by teacher feedback, improved writing performance in the first group and which writing aspects improved with this sequence. The first author taught the classes with two raters scoring the tests and assignments. Moreover, the Null hypothesis (H0) was tested in this study: "There is no significant effect of the use of Grammarly followed by teacher feedback in the writing classes of Indonesian undergraduate EFL students."

3.5. Data analysis

This section presented data analysis of the student's scores before and after the study and data analysis of the questionnaire from two classes.

3.5.1. Data analysis from the students' tests

In this study, to measure the effectiveness of Grammarly and teacher feedback, researchers calculated the Gain Score (g) followed by One-way ANCOVA, and a t-test was employed, too. One-Way ANCOVA was employed to compare two or more groups while controlling for a continuous covariate that may affect the dependent variable (Howell, 2012). Howell adds that it is a continuation of One-Way ANCOVA that eliminates a covariate's impact before examining group differences.

Next, this study was conducted to measure the extent to which Grammarly and teacher feedback affect the writing skills of Indonesian EFL undergraduate students' effectiveness of Grammarly and teacher feedback and a t-test was conducted. A t-test was used to compare the first and second experimental groups for each indicator individually and as a group (Ghufron & Rosyida, 2018). First, the pre-and post-test results were analyzed using a paired-sample t-test to see how the AWE programs affected their writing ability. Second, an independent samples t-test was used to determine whether there was a difference in using Grammarly, followed by teacher feedback in this study.

3.5.2. Data analysis of the questionnaire

In terms of quantitative data, we described students' responses towards using Grammarly and teacher feedback on their writing quality to answer the third research question. Descriptive statistics were used to answer the third research question. It aims at knowing whether students have positive responses to using Grammarly and teacher feedback. Therefore, SPSS version 20 was used to calculate the data.

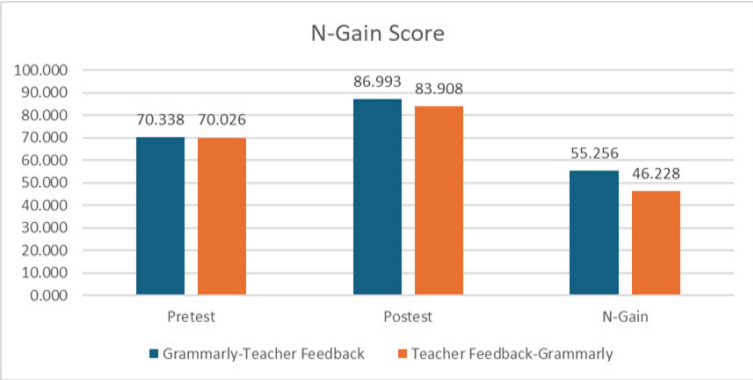
4. RESULTS

4.1. Effects of Grammarly and teacher feedback on the writing skills of Indonesian undergraduate EFL students

4.1.1. Gain score

In this study, after calculating and getting a mean score, the gain score from the first and second experimental groups was calculated, and the result is shown in Figure 2.

FIGURE 2. N-Gain score (first & second experimental groups)



In this study, as shown in Figure 2, the N-Gain score in the first experimental group was more significant than in the second experimental group. The N-Gain score in the first experimental group was 55.256 with moderate criteria, while the second group's N-Gain score was 46.228 with moderate criteria. From these data, it can be seen that there was an average increase in students' argumentative texts using Grammarly as automated feedback in the writing classes from both the tested classes. In sum, using Grammarly as automated feedback was more effective in the first experimental group than in the second group.

Then, based on the Gain score from the pre-test- and the post-test results, the mean in the first experimental group was more significant than in the second experimental group. Next, the null hypothesis in this study that the use of Grammarly and teacher feedback affects the writing skills of Indonesian undergraduate EFL students cannot be rejected. Therefore, it can be concluded that using Grammarly, followed by teacher feedback in writing classes of Indonesian undergraduate EFL students, was effective in this study.

4.1.2. One-Way ANCOVA

In this study, there were two hypotheses: Ho= There is no significant effect of using Grammarly followed by teacher feedback in the writing classes of Indonesian undergraduate EFL students. Then, for Ha=, there is a significant effect of using Grammarly followed by teacher feedback in the writing classes of Indonesian undergraduate EFL students. Based on the SPSS table output, One-Way ANCOVA has a sig value of 0.000 < 0.05. As a result, Ho was eliminated, and Ha was retained, resulting in differences in learning outcomes when using Gramarly applications and receiving teacher feedback.

TABLE 2. Tests of Between-Subjects Effects

Dependent Variable: Posttest					
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Corrected Model	484.145[a]	2	242.072	19.412	.000
Intercept	1865.755	1	1865.755	149.614	.000
Pretest	298.729	1	298.729	23.955	.000
Group	172.525	1	172.525	13.835	.000
Error	935.287	75	12.470		
Total	571482.140	78			
Corrected Total	**1419.432**	**77**			

a. R Squared = .318 (Adjusted R Squared = .300)

An analysis of covariance (ANCOVA) was performed using the first test score as a covariate to see if the means of the second test scores for the two groups differed statistically significantly. The significance level was established at p = 0.05 for all analyses. The findings confirmed the validity of the normality assumption, as shown in Table 2.

TABLE 3. Test of Normality

	Group	**Kolmogorov-Smirnov[2]**			**Shapiro-Wilk**		
		Statistic	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Residual for Posttest	Grammarly-Teacher Feedback	.122	40	.139	.968	40	.311
	Teacher Feedback -Grammarly	.096	38	.200*	.989	38	.960

Results of the normality test obtained a significance of 0.311 and 0.960, and both classes had a significance greater than 0.05. Leven's test for homogeneity of variance shows that the assumption of equality of variance is not violated; it is shown in table 3 with a significance value of 0.221> 0.05.

The sig value is determined by examining the SPSS output table of the One-Way ANCOVA test in table 4. 0.05 is less than 0.000. Consequently, there is a distinction between learning through the Grammarly application followed by teacher feedback and receiving teacher feedback followed by using the Grammarly application.

TABLE 4. Levene's Test of Equality of Error

Variances			
Dependent Variable: Posttest			
F	**df1**	**df2**	**Sig.**
1.521	1	76	.221

TABLE 5. Tests of Between-Subjects Effects

Dependent Variable: Posttest					
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Group	172.525	1	172.525	13.835	.000

TABLE 6. Tests of Group

Dependent Variable: Posttest				
Group	**Mean**	**Std. Error**	**95% Confidence Interval**	
			Lower Bound	**Upper Bound**
Grammarly-Teacher Feedback	86.940[a]	.558	85.827	88.052
Teacher Feedback-Grammarly	83.963[a]	.573	82.822	85.105

Based on the results of the descriptive analysis in table 5, the adjusted mean value in the class with Grammarly treatment followed by teacher feedback is 86.940, and in the group with teacher feedback treatment first, then using the Grammarly application gets a mean of 83.963. Thus, the conclusion of the class with Grammarly first is higher than the teacher's feedback, so the learning treatment with Grammarly has a significant effect.

4.1.3. The t-Test

Grammarly and teacher feedback were two types of feedback given to the students in this study. On the one hand, Grammarly deals with three aspects of writing: Diction, language use, and mechanics. On the other hand, teacher feedback may cover the five writing aspects: Content, Organization, Diction, Language Use, and Mechanics. These feedback types were given in the first and second experimental groups.

This study's t-test determined if the two groups wrote similarly before answering the first research question. Pre- and post-test writing scores were 1-5. A significant difference in post-test scores was observed between participants who favored Grammarly as automated feedback (M=4.0137, SD=.46362) and the experimental group (M=4.2310, SD=.61832) (Sig=.008<.05). Indonesian undergraduate EFL students benefit more from Grammarly first followed by teacher feedback in writing classes. Grammarly worked, and teacher feedback in Indonesian undergraduate EFL students' writing classes rejected H0 and accepted H1.

Indonesian undergraduate EFL students have been proven more effective than the teacher feedback first followed by Grammarly. Therefore, H_0 was rejected, and H1 was accepted because Grammarly was effective, followed by teacher feedback in the writing classes of Indonesian undergraduate EFL students.

4.2. The extent to which Grammarly and teacher feedback affect the writing skills of Indonesian EFL undergraduate students

Grammarly and teacher feedback were two types of feedback given to the students in this study. On the one hand, Grammarly deals with three aspects of writing: Diction, language use, and mechanics. On the other hand, teacher feedback may cover the five writing aspects: Content, Organization, Diction, Language Use, and Mechanics. These feedback types were given in the first and second experimental groups.

As shown in Table 2, based on the data processing results related to teacher feedback in terms of content, the mean for the first experimental group was 27.08, which is better than the second experimental group, which is 22.78. The data were normally distributed and homogeneous. For the t-test, the calculated t value is 9.189>2.040, indicating an effect between the first and second experimental groups with a significant effect of 0.3%.

TABLE 7. The summary feedback from Grammarly and teacher feedback

Items	Group	N	Mean Score	Normality Test	Homogeneity Test	t-test	t-table (α= 0.05)	Conclusion
Content	First Experimental Group	40	27.08	Normal	Homogenous	9.189	2.040	t-test>t-table accepted, Ho is rejected
	Second Experimental Group	38	22.78	Normal	Homogenous			
Organization	First Experimental Group	40	17.1842	Normal	Homogenous	8.369		t-test>t-table accepted, Ho is rejected
	Second Experimental Group	38	13.55	Normal	Homogenous			
Diction	First Experimental Group	40	18.2632	Normal	Homogenous	11.434		t-test>t-table accepted, Ho is rejected
	Second Experimental Group	38	16.82	Normal	Homogenous			
Language Use	First Experimental Group	40	19.6645	Normal	Homogenous	3.033		t-test>t-table accepted, Ho is rejected
	Second Experimental Group	38	19.50	Normal	Homogenous			
Mechanics	First Experimental Group	40	3.99	Normal	Homogenous	3.363		t-test>t-table accepted, Ho is wrejected
	Second Experimental Group	38	3.9539	Normal	Homogenous			

TABLE 8. The result of the t-test (First and Second experimental group)

Group	N	Mean	Normality	Homogeneity	t-test	t-table
First Experimental Group	40	103.87	Normal	Homogenous	1.001	1.696
Second Experimental Group	38	99.48	Normal			

Next, as shown in Table 3, several assessments in the two experimental classes support Grammarly and teacher feedback in the writing course. The pre-test and post-test results showed that the tools can improve students' writing. Data met normality and homogeneity criteria because the t-test result was 1.001, less than the t-table value of 1.696. This confirms that Grammarly and teacher feedback improve student writing.

4.3. Responses of Indonesian undergraduate EFL students to Grammarly and teacher feedback

The online questionnaire was sent to eighty subjects from two classes. However, only 78 participants completed the online questionnaire given through Google Forms, making it easier for the students to read and respond to the statements.

The students' responses to Grammarly in the writing classes were generally positive. Moreover, this questionnaire consisted of 25 statements divided into three parts: The Effectiveness of Grammarly, Grammarly as an Instructional Tool to Help with Writing from Sources, and the Frequency of Using Grammarly, as shown in Figure 2.

Figure 2 shows that statements 1, 2, and 12 had the highest score from the first part of the experimental group questionnaire. Statement 1 was that "Grammarly's feedback positively affects my writing text." Statement 2 was "Grammarly boosted my confidence as I handed in the length of argumentative texts," and Statement 12 was "Grammarly is a user-friendly program." These three statements had the highest percentage of respondents strongly agreeing (92.5%).

Then, the result of the questionnaire from the first experimental group (Second Part: Grammarly as an Instructional Tool to help with Writing from Sources) is shown in Figure 3.

FIGURE 3. Questionnaire from the first experimental group (First part: The effectiveness of Grammarly)

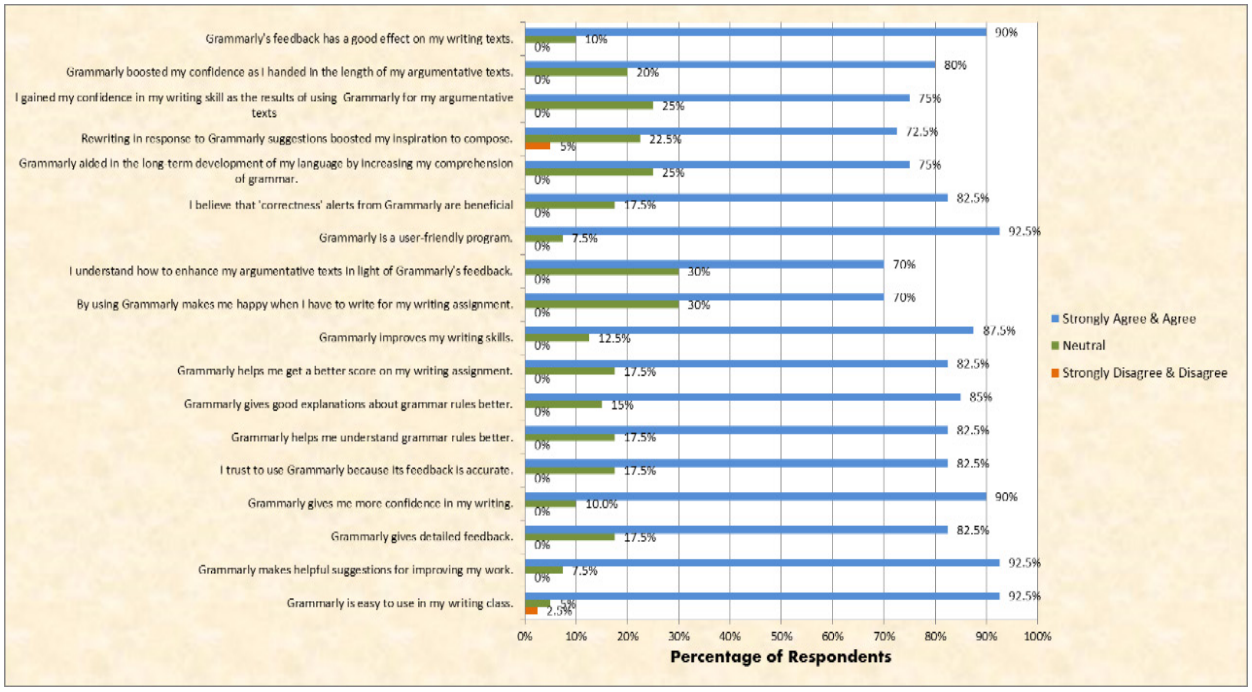


FIGURE 4. Questionnaire from the first experimental group (Second part: Grammarly as an instructional tool to help with writing from sources)

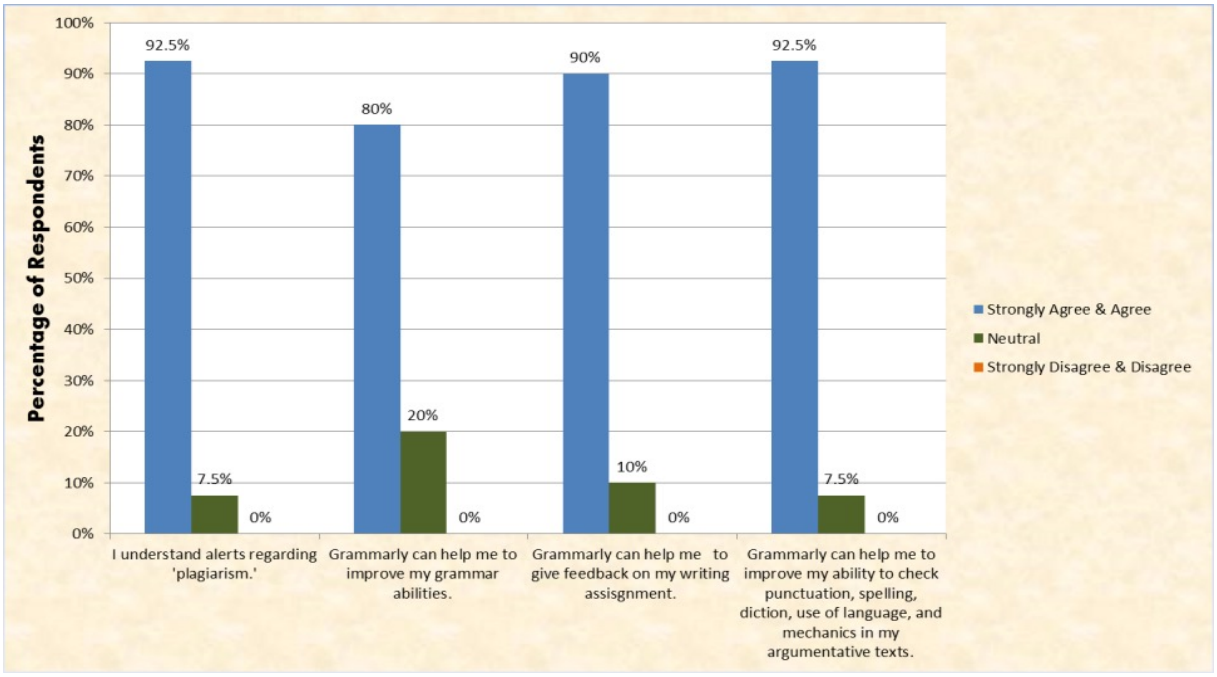
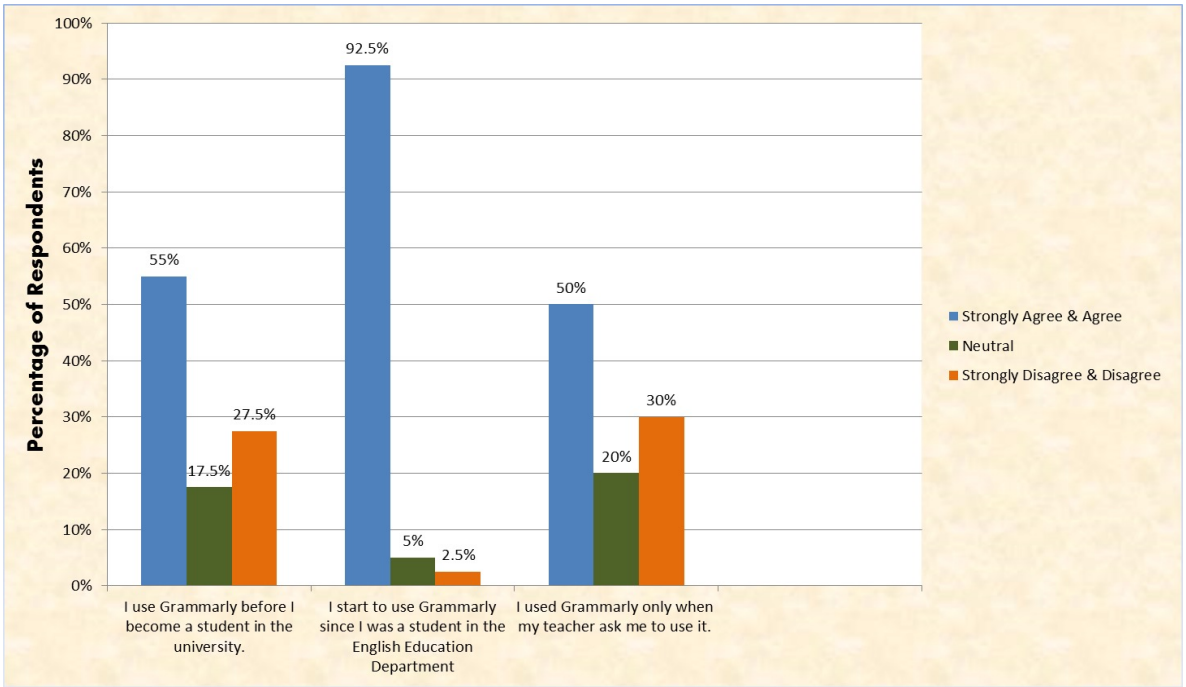


FIGURE 5. Questionnaire from the first experimental group (Third part: frequency of using Grammarly)



As shown in Figure 5, Statement 24 reached the highest score from the third part of the first experimental group with a questionnaire: "I started to use Grammarly since I was a student in the English Education Department."

5. DISCUSSION

5.1. Effects of Grammarly and teacher feedback on the writing skills of Indonesian undergraduate EFL students

This study divided the tests into two sections: pre-test and post-test. The pre-test showed that first-year students and those who had completed semester one were balanced, starting both groups at the same level. After testing basic EFL writing skills, students were assigned a topic. For the post-test, the samples were balanced differently, so the groups were treated separately. The first experimental group used Grammarly for the first two writing tasks, followed by teacher feedback for the last two. Students received Grammarly's corrected and original reports, which included revision comments for easy evaluation of their writing.

This study examined how Grammarly affected Indonesian EFL students. The first experimental group performed better than the second. The first experimental group scored 68.61% on N-Gain effectiveness interpretation, indicating success. The second experimental group's average N-Gain score was 46.675, or 46.68 percent, indicating an inadequate performance. Indonesian college students learning English as a second language performed better in the first group that used Grammarly first and then teacher feedback.

This study could extend previous research on Grammarly and teacher feedback. Grammarly and teacher feedback showed they successfully integrated input into students' revisions, and their increased post-test scores showed the students used feedback and improved their writing, demonstrating its pedagogical potential for students (Thi & Nikolov, 2022). Nova (2018) said that Grammarly's feedback helps them learn better. O'Neill and Russell (2019) stated that Grammarly provides more feedback faster than traditional methods. The findings suggest using machine feedback in writing instruction to supplement teacher feedback. The findings suggest using machine feedback in writing instruction to supplement teacher feedback. The findings suggest using machine feedback in writing instruction to supplement teacher feedback. Ghufron and Rosyida (2018) also found that students who use Grammarly make fewer errors than those who use a teacher (indirect corrective feedback). Moreover, Grammarly's user-friendliness and accessibility make it simple for students to employ, resulting in favorable reviews and enhanced writing practices (Raskova, 2023). In sum, Grammarly's use in this study was effective, particularly in students' argumentative writing skills for Indonesian EFL undergraduates.

Previous research has demonstrated that Grammarly's feedback feature assists students with writing improvement and error identification (Qassemzadeh & Soleimani, 2016). Daniels and Leslie (2016) also reported that Grammarly is language learning software that can evaluate EFL writing and improve language skills. This is also consistent with our earlier findings, which demonstrated that using electronic feedback such as Grammarly software assists students in overcoming their apprehension of grammar (Saadi & Saadat, 2015). Then Grammarly is a helpful instrument for learning English, providing many tools and challenging exercises that keep students involved (Galingging et al., 2023). Moreover, Grammarly offers valuable features such as grammar checker, punctuation, and spelling, helping students save time and develop confidence in academic writing (Faisal & Carabella, 2023). In sum, Grammarly's feedback feature facilitates students improve their writing skills by recognizing mistakes by providing an abundance of tools and challenging exercises that ensure they remain engaged.

5.2. The extent to which Grammarly and teacher feedback affect the writing skills of Indonesian EFL undergraduate students

The second question of the study investigates to what extent Grammarly affects the writing skills of Indonesian undergraduate EFL students. The investigation of using Grammarly on students' texts was calculated using Diction, Language Use, and Mechanics. Then, teacher feedback was on Content, Organization, Mechanics, Language Use, and Diction. This study found that Grammarly software has more effect on mechanics (spelling and punctuation), language use (grammar), and diction but less on the content and organization. However, teacher feedback has more effect on content and organization but less on mechanics, language use, and diction.

Several reports have shown that students who use Grammarly to evaluate their work have significantly fewer errors than students who use the teacher's evaluation, and writing teachers could use it regularly or encourage students to use it independently (Ghufron & Rosyida, 2018; Thi & Nikolov, 2022). Ghufron and Rosyida add that Grammarly reduces diction, language use, and writing mechanics errors. Then, Dizon and Gayed (2021) stated that based on the descriptive statistics and t-test results, they made fewer grammatical

errors when students used Grammarly to help them write in L2. Next, Grammarly emphasizes the significance of technology and self-learning in modern learning by enabling writers to track learning objectives and engagement (Wardatin et al., 2022). Moreover, Grammarly has the potential to be a helpful tool for facilitating students' learning and assessment of source-based writing techniques (Dong & Shi, 2021). In sum, in this study, using Grammarly significantly affects the mechanics of the argumentative texts written by students rather than the diction and language use.

In the first experimental group, Grammarly was used to evaluate and correct students' writing for the first two texts, with students trained to use it independently. Teachers then provided feedback on the third and fourth texts. Grammarly reports, submitted alongside the corrected documents, included comments that helped teachers assess revisions. Students were evaluated on writing mechanics (spelling, punctuation, grammar) and diction.

For the third and fourth assignments, teachers reviewed student work, reading it aloud and taking notes on content, organization, diction, mechanics, and language use. Corrections were made and returned via Google Drive. Grammarly effectively reduced mechanics, diction, and language errors, identifying mistakes such as missing spaces and punctuation while offering correction suggestions. This aligns with research showing Grammarly aids EFL writing evaluation. It aligns with Daniels and Leslie (2016) that Grammarly is a software application that can aid language learning, particularly when evaluating EFL writing. Grammarly's feedback feature also helps students identify errors and improve their writing (Qassemzadeh & Soleimani, 2016; Thi & Nikolov, 2022).

Grammarly has minimal impact on writing organization and content, as it cannot assess topic relevance or paragraph coherence. While it detects sentence movement, it overlooks logical flow. The tool prioritizes mechanical accuracy—spelling, punctuation, and grammar—over argumentation and coherence (Muhammad, 2024; Resiana et al., 2024). It may also flag contextually appropriate sentences as incorrect, emphasizing form over substance. This reliance on technical correctness can limit critical thinking and complex writing, potentially hindering academic success.

Next, the limitation of Grammarly is that it cannot give students personalized feedback that meets their learning needs (Hasby et al., 2024). Therefore, without tailored guidance, students may receive the same feedback regardless of their writing strengths or weaknesses (Ding & Zou, 2024). Moreover, Grammarly may not help writers improve due to its lack of personalized support (Ebadi et al., 2022). In conclusion, Grammarly can provide general corrections but not customized feedback to help students improve their writing.

Regarding the other issues and limitations of using Grammarly in writing classes. Some students initially do not have computers, tablets, or a stable internet connection to use Grammarly (Sanderson & Stephens, 2023). Without devices or reliable internet, students cannot use the tool to correct their grammar, making automated feedback less useful. Second, students may not be comfortable using Grammarly or other digital tools. Software novices may struggle to use Grammarly's features, which may frustrate or prevent them from using it (Giray, 2024). Third, it is related to Grammarly in the form of free and premium limits. Grammarly's free version may not give students as much feedback as the paid version, hindering their writing improvement (Setyowati et al., 2024). Students who cannot afford the premium version may miss out on advanced features, creating disparities. While Grammarly supports writing, its accessibility, usability, and cost barriers

must be addressed. However, although some previous studies have shown the limitations of Grammarly, related to the result, this study found that it is more effective to begin with Grammarly and then provide feedback to the teacher rather than providing feedback to the teacher first. This research implies that. This fostered greater autonomy among students and motivated them to verify their assignments. Grammarly affects mechanics, diction, language use, and teacher feedback concerning content and organization.

Concerning the second research question, this study has shown that the present results are significant in at least two crucial respects. The findings in this study align with Barrot (2020), who states that Grammarly is an effective tool for students to use in their writing classes because it helps detect mechanical errors. Next, the findings in this study support the previous study from Fitria (2022) that Grammarly automatically verifies typed work based on several factors and reveals grammatical and mechanical student writing errors. In this study, Grammarly has more effect on diction, language use, and mechanics but less on content and organization. On the other hand, teacher feedback has more effect on content and organization but less effect on diction, mechanics, and language use.

5.3. Responses to Grammarly and Teacher Feedback

Regarding the effectiveness of Grammarly, most students found the suggestions helpful in improving their papers, and half of the students believed it aided them in achieving a higher grade. In the first section of the survey, most students agreed that Grammarly was a user-friendly program that helped them improve their writing skills. This study's findings indicate that students have positive attitudes toward using AWE tools to improve writing skills. Overall, students found Grammarly to be practical and valuable. More than 85% of students rated practical and helpful factors as positive. It is in line with Perdana et al. (2021) that the Grammarly app is helpful for many linguistic issues in academic writing and the outside academic world.

The findings of this study coincide with the findings of Miranty et al. (2023), who state that Grammarly is perceived as a helpful AWE tool by students across all year levels because they know the need for proofreading services. Next, O' Neil & Russel (2019) reported that Grammarly students scored higher on 9 out of 15 survey items and were happier with the grammar advice they received than non-Grammarly students. Research by Bailey and Lee (2020) found that students using Grammarly as automated feedback can increase their confidence and save time, incrementally improving writing compositions when writing in a second language because of fewer accuracy mistakes. Moreover, undergraduate EFL students responded positively regarding AWE tools' potential advantages, such as Grammarly, in enhancing their writing skills (Miranty et al., 2023).

Concerning the third research question, based on the result from the questionnaire, there were positive responses from learners on using Grammarly as automated feedback instead of teacher feedback for submission in their writing course. Grammarly positively influenced the learners, gave feedback, and corrected their mistakes before submission, thus saving instructor effort and developing self-assisted learning styles among EFL learners.

6. CONCLUSIONS

This study examined using Grammarly and teacher feedback in undergraduate EFL writing classes in Indonesia. These two types of feedback, namely Grammarly and teacher feedback, were applied because Grammarly could reduce the teacher's time to check student assignments, make students independent learners, and give immediate writing feedback. However, Grammarly cannot replace teachers because ELT classrooms require student-teacher interaction. Grammarly in EFL writing reduces errors in mechanics (spelling and punctuation), language use (grammar), and diction (vocabulary), but not content and organization. Teacher feedback affects content and organization more than diction, language, and mechanics. The teacher spotted topic-content gaps quickly. Therefore, when the paragraph lacks coherence, the teacher may be sensitive. Moreover, incorporating Grammarly and teacher feedback into students' revisions showed that students got feedback effectively, which improved their writing skills. In sum, Grammarly and teacher feedback are valuable tools for enhancing EFL writing skills in Indonesian undergraduate students.

The study indicates two primary implications: first, Grammarly aids large classes in conserving time by delivering instantaneous feedback on mechanics, whereas teacher feedback is crucial for enhancing content and organization. Second, the research indicated that employing Grammarly before receiving teacher feedback proved more effective. The study's dependence on closed-ended questions constrained its breadth, necessitating future research to include open-ended questions for enhanced understanding. Subsequent research should investigate various contexts, the influence of automated feedback on self-editing, and the responses of L1 and EFL learners to both automated and instructor feedback. Further investigation is required to validate the effectiveness of automated feedback and its overall influence on writing enhancement.

Declaration of conflicting interest

There is no conflict of interest in this work.

7. FUNDING

We want to express our sincere gratitude to the Research and Community Service Unit of Universitas Negeri Malang with contract number 19.5.1104/UN32.20.1/LT/2022 for funding our research, which was used as a basis to write this article.

8. REFERENCES

Agustin, R., & Wulandari, S. (2022). The Analysis of Grammatical Errors on Students' Essay Writing by Using Grammarly. *Jurnal Pendidikan Bahasa Inggris Proficiency, 4*(1), 39–46. https://doi.org/10.32503/proficiency.v4i1.2247

Albesher, K. B. (2022). Teachers' Views on Using the Writing Process Approach to Improve ESL Learners' Writing Skills. *International TESOL & Technology Journal (ITTJ), 17*(2), 76–95. https://connect.academics.education/index.php/itj/issue/view/31

Alharbi, W. (2023). AI in the Foreign Language Classroom: A Pedagogical Overview of Automated Writing Assistance Tools. Education Research International. https://doi.org/10.1155/2023/4253331

Aljameel, I. H. (2022). Computer-Assisted Language Learning in Saudi Arabia: Past, Present, and Future. *International Education Studies, 15*(4), 95. https://doi.org/10.5539/ies.v15n4p95

Almusharraf, N., & Alotaibi, H. (2023). An error-analysis study from an EFL writing context: Human and Automated Essay Scoring Approaches. *Technology, Knowledge and Learning, 28*(3), 1015–1031. https://doi.org/10.1007/s10758-022-09592-z

Alsmari, N. A. (2019). Fostering EFL Students' Paragraph Writing Using Edmodo. *English Language Teaching, 12*(10), 44. https://doi.org/10.5539/elt.v12n10p44

Anastasia, W., Murtisari, E., & Isharyanti, N. (2024). Less-proficient EFL students' use of Grammarly in writing: Behavior, cognition, and affect. *The JALT CALL Journal, 20*(1), 1–25. https://doi.org/10.29140/jaltcall.v20n1.1089

Andina, D. M., Cahyono, B. Y., & Widiati, U. (2020). How English Foreign Language Students' Autonomy and Digital Competence Relate to Their Writing Achievement. *Tadris: Jurnal Keguruan Dan Ilmu Tarbiyah, 5*(1), 77–86. https://doi.org/10.24042/tadris.v5i1.5760

Ariyanto, M. S. A., Mukminatien, N., & Tresnadewi, S. (2021). College Students' Perceptions of an Automated Writing Evaluation as a Supplementary Feedback Tool in a Writing Class. *Jurnal Ilmu Pendidikan (JIP), 27*(1), 41–51. https://doi.org/10.17977/um048v27i1p41-51

Arredondo, L. L. A., & Laurens, L. (2023). Metaversity: Beyond Emerging Educational Technology. *Sustainability, 15*(22), 15844. https://doi.org/10.3390/su152215844

Awada, G. M., & Diab, N. M. (2023). Effect of online peer review versus face-to-Face peer review on argumentative writing achievement of EFL learners. *Computer Assisted Language Learning, 36*(1–2), 238–256. https://doi.org/10.1080/09588221.2021.1912104.

Ayan, A. D., & ErdemiR, N. (2023). EFL Teachers' Perceptions of Automated Written Correctiv Feedback and Grammarly. *Ahmet Keleşoğlu Eğitim Fakültesi Dergisi, 5*(3), 1183-1198. https://doi.org/10.38151/akef.2023.106

Backman, Y., Reznitskaya, A., Gardelli, V., & Wilkinson, I. A. G. (2023). Beyond Structure: Using the Rational Force Model to Assess Argumentative Writing. *Written Communication, 40*(2), 555–585. https://doi.org/10.1177/07410883221148664

Bai, B., Wang, J., & Zhou, H. (2022). An Intervention Study to Improve Primary School Students' self-regulated strategy use in English Writing through e-learning in Hong Kong. *Computer Assisted Language Learning, 35*(9), 2265–2290. https://doi.org/10.1080/09588221.2020.1871030

Bai, L., & Hu, G. (2017). In the face of fallible AWE feedback: How do students respond? *Educational Psychology, 37*(1), 67–81. https://doi.org/10.1080/01443410.2016.1223275

Bailey, D., & Lee, A. R. (2020). An Exploratory Study of Grammarly in the Language Learning Context: An Analysis of Test-Based, Textbook-Based and Facebook Corpora. *TESOL International Journal, 15*(2), 4–27. https://www.tesol-international-journal.com/volume-15-issue-2-2020/

Balta, E. E. (2018). The Relationships Among Writing Skills, Writing Anxiety and Metacognitive Awareness. *JEL, 7*(3), 233-241. https://doi.org/10.5539/jel.v7n3p233

Barrot, J. S. (2020). Integrating Technology into ESL/EFL Writing through Grammarly. *RELC Journal*, 1–5. https://doi.org/10.1177/0033688220966632

Barrot, J. S. (2023). Trends in automated writing evaluation systems research for teaching, learning, and assessment: A bibliometric analysis. *Education and Information Technologies,* 7155–7179. https://doi.org/10.1007/s10639-023-12083-y

Boud, D., & Dawson, P. (2023). What feedback literate teachers do: An empirically-derived competency framework. *Assessment & Evaluation in Higher Education, 48*(2), 157–171. https://doi.org/10.1080/02602938.2021.191092

Bulatović, V. V., Mirović, I., & Kaurin, T. (2024). Analyzing grammarly software for corrective feedback: Teacher's perspective on affordances, limitations and implementation. *Focus on ELT Journal, 6*(1), 74–86. https://doi.org/10.14744/felt.6.1.6

Cavaleri, M., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning, 10*(1), 223–236. https://journal.aall.org.au/index.php/jall/article/view/393

Daniels, P., & Leslie, D. (2016). Grammar software ready for EFL writers? *OnCue Journal, 9*(4), 391–401. https://jaltcue.org/journal_9.4

Dewi, U. (2023). Grammarly as Automated Writing Evaluation: Its Effectiveness from EFL Students' Perceptions. *Lingua Cultura, 16*(2), 155–161. https://doi.org/10.21512/lc.v16i2.8315

Ding, L., & Zou, D. (2024). Automated writing evaluation systems: A systematic review of Grammarly, Pigai, and Criterion with a perspective on future directions in the age of generative artificial intelligence. *Education and Information Technologies, 29*, 14151–14203. https://doi.org/10.1007/s10639-023-12402-3

Dizon, G., & Gayed, J. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *The JALT CALL Journal, 17*(2), 74–92. https://doi.org/10.29140/jaltcall.v17n2.336.

Dong, Y., & Shi, L. (2021). Using Grammarly to support students' source-based writing practices. *Assessing Writing, 50*, 100564. https://doi.org/10.1016/j.asw.2021.100564

Ebadi, S., Gholami, M., & Vakili. (2022). Investigating the Effects of Using Grammarly in EFL Writing: The Case of Articles. *Computers in the Schools, 40*(1), 85–105. https://doi.org/10.1080/07380569.2022.2150067

Ebyary, K. E., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *IJES, 10*(2), 121–142. https://doi.org/10.6018/ijes/2010/2/119231.

Elumalai, K. V. (2019). Teacher Constructed Corrective Feedback Enhancing Students Writing Skills in EFL Classroom. *Advances in Language and Literary Studies, 10*(5), 103. https://doi.org/10.7575/aiac.alls.v.10n.5p.103

Estacio, D. L., Valencia, N. M., & Abdala, M. P. L. T. (2022). Development and Evaluation of Academic Learning Modules by the Faculty of Architecture and Fine Arts of Bulacan State University Using Grammarly Software by the Library Service. *International Journal of Research Publications, 97*(1), 335–344. https://doi.org/10.47119/IJRP100971320223002

Fahmi, M. A., & Cahyono, B. Y. (2021). EFL students' perception on the use of Grammarly and teacher feedback. *JEES (Journal of English Educators Society), 6*(1), 18–25. https://doi.org/doi:https://doi.org/10.21070/jees.v6i1.849.

Faisal, F., & Carabella, P. A. (2023). Utilizing Grammarly in an Academic Writing Process: Higher-Education Students' Perceived Views. *Journal of English Language Teaching and Linguistics, 8*(1), 23-42. https://doi.org/10.21462/jeltl.v8i1.1006

Fitria, T. N. (2021). Grammarly' As A Teachers' Alternative in Evaluation Non-EFL Students Writings. *Leksema: Jurnal Bahasa Dan Sastra, 6*(2), 141–152. https://doi.org/10.22515/ljbs.v6i2.3957

Fitria, T. N. (2022). Identifying Grammatical and Mechanical Errors of Students' Writing: Using "Grammarly" as an Online Assessment. *Lingua Didaktika: Jurnal Bahasa Dan Pembelajaran Bahasa, 16*(2), 169–184. https://doi.org/10.24036/ld.v16i2.116824

Fitriana, K., & Nurazni, L. (2021). Exploring English Department Students' Perceptions on Using Grammarly to Check the Grammar in their Writing. *Journal of English Teaching, 8*(1), 15–25. https://doi.org/10.33541/jet.v8i1.3044

Fitriana, K., & Nurazni, L. (2022). Exploring Students' Perception of Using Grammarly to Check Grammar in Their Writing. *JET (Journal of English Teaching), 8*(1), 15–25. https://doi.org/10.33541/jet.v8i1.3044

Galingging, C. K., Sipayung, K. T., Silitonga, H., & Pardede, S. (2023). The Effectiveness Of Grammarly Application On Writing Descriptive Text Tenth Grade SMA Negeri 1 Lau Baleng. *Journal on Education, 06*(01), 2891–2904. https://jonedu.org/index.php/joe

Ghanbari, N., & Salari, M. (2022). Problematizing Argumentative Writing in an Iranian EFL Undergraduate Context. Frontiers in Psychology, 13, 862400. https://doi.org/10.3389/fpsyg.2022.862400

Ghufron, M. A., & Rosyida, F. (2018). The Role of Grammarly in Assessing English as a Foreign Language (EFL) Writing. *Lingua Cultura, 12*(4), 395. https://doi.org/10.21512/lc.v12i4.4582.

Giray, L. (2024). "Don't Let Grammarly Overwrite Your Style and Voice:" Writers' Advice on Using Grammarly in Writing. *Internet Reference Services Quarterly, 28*(3), 293–303. https://doi.org/10.1080/10875301.2024.2344762

Gozali, I., Wijaya, A. R. T., Lie, A., Cahyono, B. Y., & Suryati, N. (2024). Leveraging the potential of ChatGPT as an automated writing evaluation (AWE) tool: Students' feedback literacy development and AWE tools integration framework. *The JALT CALL Journal, 20*(1), 1–22. https://doi.org/10.29140/jaltcall.v20n1.1200

Grammarly. (2019). *About Grammarly*. Retrieved from https://app.grammarly.com.

Hasby, N., Perdana, I., & Kodriyah, L. (2024). A Systematic Literature Review on Grammarly in English Studies: Advantages that Dazzle and Disadvantages that Fazzle. *AJELP: The Asian Journal of English Language & Pedagogy, 12*(2), 139–154.

Heron, M., Medland, E., Winstone, N., & Pitt, E. (2023). Developing the relational in teacher feedback literacy: Exploring feedback talk. *Assessment & Evaluation in Higher Education, 48*(2), 172–185.https://doi.org/10.1080/02602938.2021.1932735

Hockly, N. (2019). Automated Writing Evaluation. *ELT Journal, 73*(1), 82–88. https://doi.org/10.1093/elt/ccy044

Howell, D. C. (2012). *Statistical Methods for Psychology (8th ed.)*. Cengage Learning.

Inayah, T. M., & Apoko, T. W. (2024). Exploring Students' Perspectives on the Use of Grammarly in Writing Analytical Exposition Text. *JLE: Journal of Literate of English Education Study Program, 5*(1), 73–83. https://doi.org/10.47435/jle.v5i1.2802

Jacobs, H. L., Zingkraf, S. A., Wormuth, D. R., Hearfiel, V. F., & Hughey, J. B. (1981). *Testing ESL Composition: A Practical Approach. English Composition Program Rowley*. MA: Newbury House. https://doi.org/10.1177/026553228400100210

Jomaa, N., & Jibroo, H. (2024). Corrective Feedback of Grammarly in Enhancing L2 Writing by EFL Kurdish Students. *Bulletin of Advanced English Studies, 9*(1), 1–15. https://doi.org/10.31559/BAES2024.9.1.1

Joo, M. (2021). The Influence of Users' Satisfaction With AWE on English Learning Achievement through Self-Efficacy: Using PLS-SEM. *Journal of Digital Convergence, 19*(9), 1–8. https://doi.org/10.14400/JDC.2021.19.9.001

Jumariati, J., & Sulistyo, G. (2017). Problem-Based Writing Instruction: Its Effect on Students' Skills in Argumentative Writing. *Arab World English Journal, 8*(2), 87–100. https://doi.org/10.24093/awej/vol8no2.6

Karyuatry, L., Rizqan, M. D., & Darayani, N. A. (2018). Grammarly as a Tool to Improve Students' Writing Quality: Free Online-Proofreader across the Boundaries. *JSSH (Jurnal Sains Sosial Dan Humaniora), 2*(1), 83–89. https://doi.org/10.30595/jssh.v2i1.2297

Kleemola, K., Hyytinen, H., & Toom, A. (2022). The Challenge of Position-Taking in Novice Higher Education Students' Argumentative Writing. *Front. Educ, 7*. https://doi.org/10.3389/feduc.2022.885987

Koltovskaia, S. (2023). Postsecondary L2 writing teachers' use and perceptions of Grammarly as a complement to their feedback. *ReCALL, 35*(3), 290–304. https://doi.org/10.1017/S0958344022000179

Khushk, M., Masroor, H., & Naeem, A. (2024). Evolution and Use of Grammarly in English Language Teaching and Learning in The Classroom Context. Al – Aijaz, *Research Journal of Islamic Studies & Humanities, VIII*(2), 72–78.

Kurnia, A. (2022). EFL Students' Problems In Dealing With Teacher Written Feedback. *Jurnal Pendidikan Indonesia, 3*(05), 485–492. https://doi.org/10.36418/japendi.v3i05.1143

Landrieu, Y., De Smedt, F., Van Keer, H., & De Wever, B. (2022). Assessing the Quality of Argumentative Texts: Examining the General Agreement Between Different Rating Procedures and Exploring Inferences of (Dis)agreement Cases. *Frontiers in education, 7*, 784261 https://doi.org/10.3389/feduc.2022.784261

Landrieu, Y., De Smedt, F., Van Keer, H., & De Wever, B. (2023). Argumentation in collaboration: The impact of explicit instruction and collaborative writing on secondary school students' argumentative writing. *Reading and Writing, 37*, 1407–1434. https://doi.org/10.1007/s11145-023-10439-x

Lei, J. I. (2020). An AWE-Based Diagnosis of L2 English Learners' Written Errors. *English Language Teaching, 13*(10), 111–119. https://doi.org/10.5539/elt.v13n10p111

Li, R. (2023). Still a fallible tool? Revisiting effects of automated writing evaluation from activity theory perspective. British Journal of Educational Technology, 54(3), 773–789. https://doi.org/10.1111/bjet.13294

Liao, H. C. (2016). Using automated writing evaluation to reduce grammar errors in writing. *ELT Journal, 70*(3), 308–319. https://doi.org/10.1093/elt/ccv058.

Liu, C. C., Liu, S. J., Hwang, G. J., Tu, Y. F., Wang, Y., & Wang, N. (2023). Engaging EFL students' critical thinking tendency and in-depth reflection in technology-based writing

contexts: A peer assessment-incorporated automatic evaluation approach. *Educ Inf Technol, 28*, 13027-13052. https://doi.org/10.1007/s10639-023-11697-6

Loncar, M., Schams, W., & Shing, L. (2023). Multiple technologies, multiple sources: Trends and analyses of the literature on technology-mediated feedback for L2 English writing. *Computer Assisted Language Learning, 36*(4). https://doi.org/10.1080/09588221.2021.1943452

Martínez, J., López-Díaz, A., & Pérez, E. (2020). Using Process in the Teaching English as a Foreign Language. *Revista Caribeña de Investigación Educativa (RECIE), 4*(1), 42–61. https://doi.org/doi: https://doi.org/10.32541/recie.2020.v4i1.pp49-61

Mirzae, F., & Shamsudin, S. (2023). Process Oriented Writing Approach by UTM Iranian Students: Difficulties and Features. *International Journal of Advanced Research in Education and Society, 5*(2), 248–260. https://doi.org/10.55057/ijares.2023.5.2.23

Miranty, D., & Widiati, U. (2021). An automated writing evaluation (AWE) in higher education. *Pegem Journal of Education and Instruction, 11*(4), 126–137. https://doi.org/10.47750/pegegog.11.04.12

Miranty, D., Widiati, U., Cahyono, B. Y., & Suzila, T. I. (2023). Automated writing evaluation tools for Indonesian undergraduate English as a foreign language students' writing. *International Journal of Evaluation and Research in Education (IJERE), 12*(3), 1705. https://doi.org/10.11591/ijere.v12i3.24958

Mohsen, M. A., & Alshahrani, A. (2019). The Effectiveness of Using A Hybrid Mode of Automated Writing Evaluation System on EFL Students' Writing. *Teaching English with Technology, 19*(1), 118–131. https://tewtjournal.org/volume-2019/issue-1/

Muhammad, G. (2024). The Effectiveness of Grammarly Features in Building Arguments in Writing Essays. *IKOMTI, 5*(3), 41–47. https://doi.org/10.35960/ikomti.v5i3.1664

Nasser, S. M. (2018). Iraqi EFL Students' Difficulties in Writing Composition: An Experimental Study (University of Baghdad). *International Journal of English Linguistics, 9*(1), 178. https://doi.org/10.5539/ijel.v9n1p178

Nova, M. (2018). Utilizing Grammarly in Evaluating Academic Writing: A Narrative Research on EFLStudents' Experience. *Premise: Journal of English Education and Applied Linguistics, 7*(1), 80–96. https://ojs.fkip.ummetro.ac.id/index.php/english/issue/view/April%20Edition2018.

O' Neil, R., & Russel, A. (2019). Grammarly: Help or hindrance? Academic Learning Advisors' perceptions of an online grammar checker. *Journal of Academic Language & Learning, 13*(1), 88–107. https://journal.aall.org.au/index.php/jall/article/view/591

Olson, C. B., Maamuujav, U., Steiss, J., & Chung, H. (2023). Examining the Impact of a Cognitive Strategies Approach on the Argument Writing of Mainstreamed English Learners in Secondary School. *Written Communication, 40*(2), 373–416. https://doi.org/10.1177/07410883221148724

ONeill, R., & Russell, A. (2019). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology (AJET), 35*(1), 42–55. https://doi.org/10.14742/ajet.3795

Özdemir, S. (2018). The Effect of Argumentative Text Pattern Teaching on Success of Constituting Argumentative Text Elements. *WJE, 8*(5), 112. https://doi.org/10.5430/wje.v8n5p112.

Palermo, C., & Wilson, J. (2020). Implementing Automated Writing Evaluation in Different Instructional Contexts: A Mixed-Methods Study. *Journal of Writing Research, 12*(1), 63–108. https://doi.org/doi: https://doi.org/10.17239/jowr-2020.12.01.04

Parra, G. L., & Calero, S. X. (2019). Automated Writing Evaluation Tools in the Improvement of the Writing Skill. *International Journal of Instruction, 12*(2), 209–226. https://doi.org/doi: https://doi.org/10.29333/iji.2019.12214a

Pearson, W. S. (2018). Written Corrective Feedback in IELTS Writing Task 2: Teachers' Priorities, Practices, and Beliefs. *The Electronic Journal for English as a Second Language, 21*(4), 1–32. https://tesl-ej.org/wordpress/issues/volume21/ej84/

Perdana, I., Manullang, S. O., & Masri, F. A. (2021). Effectiveness of Online Grammarly Application in Improving Academic Writing: Review of Experts Experience. *International Journal of Social Sciences, 4*(1), 122–130. https://doi.org/10.31295/ijss.v4n1.1444

Qassemzadeh, A., & Soleimani, H. (2016). The Impact of Feedback Provision by Grammarly Software and Teachers on Learning

Passive Structures by Iranian EFL Learners. *Theory and Practice in Language Studies, 6*(9), 1884-1894. http://dx.doi.org/10.17507/tpls.0609.23

Ranalli, J., Link, S., & Hudilainen, E. C. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *An International Journal of Experimental Educational Psychology, 47*(1), 8–25. https://doi.org/10.1080/01443410.2015.1136407

Raskova, L. (2023). Integrating Grammarly Tools to Enhance Writing Efficiency in Senior High School. *Jadila: Journal of Development and Innovation in Language and Literature Education, 3*(1), 92–106. https://doi.org/10.52690/jadila.v3i1.389

Resiana, A. T., Zamzam, A., Putera, L. J., Amrullah, & Arfah, H. (2024). The Effectiveness of Grammarly Application on the Students' Argumentative Writing Progress. *Journal of English Education Forum (JEEF), 4*(3), 153–159. https://doi.org/10.29303/jeef.v4i3.722

Roscoe, R., Wilson, J., Johnson, A. C., & Mayra, C. R. (2017). Presentation, expectations, and experience: Sources of student perceptions of automated writing evaluation. *Computers in Human Behavior, 70,* 207-221. https://doi.org/10.1016/j.chb.2016.12.076

Saadi, Z. K., & Saadat, M. (2015). EFL Learners' Writing Accuracy: Effects of Direct and Metalinguistic Electronic Feedback. TPLS, 5(10), 2053-2063. http://dx.doi.org/10.17507/tpls.0510.11

Saman, E., Gholami, M., & Vakili, S. (20233). Investigating the Effects of Using Grammarly in EFL Writing: The Case of Articles. *Computers in the Schools, 40*(1). https://doi.org/10.1080/07380569.2022.2150067.

Sanderson, I. J., & Stephens, O. S. (2023). The Frequency and Accuracy of Prompts and Suggestions Made by Free Grammarly on Thai University Students' English Writing Errors. *Chiang Mai University Journal of Humanities, 24*(3), 257–294.

Sermsook, K., Liamnimitr, J., & Pochakorn, R. (2017). An Analysis of Errors in Written English Sentences: A Case Study of Thai EFL Students. *English Language Teaching, 10*(3), 101–110. https://doi.org/10.5539/elt.v10n3p101

Setyani, E. D., Bunau, E., & Rezeki, Y. S. (2023). The Influence of Grammarly towards Indonesian EFL Students' First-Degree Thesis Writing Confidence. *Elsya: Journal of English Language Studies, 5*(1), 54–67. https://doi.org/10.31849/elsya.v5i1.6773

Setyowati, Y., Priyambudi, S., & Wijayanti, G. C. (2024). Students' Reflections on Grammarly as a Tool for Academic Writing Support: Perceived Knowledge and Challenges in Higher Education. *SCOPE: Journal of English Language Teaching, 09*(01), 577–587. https://doi.org/10.30998/scope.v9i1.24854

Shadiev, R., & Feng, Y. (2023). Using automated corrective feedback tools in language learning: A review study. *Interactive Learning Environments*, *32*(6), 2538–2566. https://doi.org/10.1080/10494820.2022.2153145

Shum, S. B., Lim, L.-A., Boud, D., Bearman, M., & Dawson, P. (2023). A comparative analysis of the skilled use of automated feedback tools through the lens of teacher feedback literacy. *International Journal of Educational Technology in Higher Education, 20*(1), 40. https://doi.org/10.1186/s41239-023-00410-9

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Stevenson, M, A. (2016). Critical Interpretative Synthesis: The Integration of Automated Writing Evaluation into Classroom Writing Instruction. *Computers and Composition, 42*(1), 1–16. http://dx.doi.org/10.1016/j.compcom.2016.05.001

Tang, J., & Rich, C. (2017). Automated Writing Evaluation in an EFL Setting: Lessons from China. *The JALT CALL Journal, 13*(2), 117–146. https://doi.org/10.29140/jaltcall.v13n2.215

Taskiran, A., & Goksel, N. (2022). Automated Feedback and Teacher Feedback: Writing Achievement in Learning English as A Foreign Language at A Distance. *Turkish Online Journal of Distance Education, 23*(2), 120–139. https://doi.org/10.17718/tojde.1096260

Teng, M. F., Wang, C., & Zhang, L. J. (2022). Assessing self-regulatory writing strategies and their predictive effects on young EFL learners' writing performance. *Assessing Writing, 51*, 100573. https://doi.org/10.1016/j.asw.2021.100573

Thi, N. K., & Nikolov, M. (2022). How Teacher and Grammarly Feedback Complement One Another in Myanmar EFL Students' Writing. *The Asia-Pacific Education Researcher, 31*(6), 767–779. https://doi.org/10.1007/s40299-021-00625-2

Thi, N. K., Nikolov, M., & Simon, K. (2022). Higher-proficiency students' engagement with and uptake of teacher and Grammarly feedback in an EFL writing course. *Innovation in Language Learning and Teaching, 17*(3), 690–705. https://doi.org/10.1080/17501229.2022.2122476

Tian, L., & Zhou, Y. (2020). Learner Engagement with Automated Feedback, peer feedback and teacher feedback in an online EFL writing context. *System, 91*. https://doi.org/doi: https://doi.org/10.1016/j.system.2020.102247

Ventayen, R. J. M., & Ventayen, C. C. O. (2018). Graduate Students' Perspective on the Usability of Grammarly® in one ASEAN State University. *Asian ESP Journal, 14*(7.2), 09–30. https://www.asian-esp-journal.com/volume-14-issue-7-2-december-2018/

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education, 194*, 104703. https://doi.org/10.1016/j.compedu.2022.104703

Wang, Z., & Chiu, M. M. (2024). Multi-discourse Modes in Student Writing: Effects of Combining Narrative and Argument Discourse Modes on Argumentative Essay Scores. *Applied Linguistics, 45*(1), 20–40. https://doi.org/10.1093/applin/amac073

Wardatin, F. N., Setiawan, S., Mustofa, A., & Nugroho, H. A. (2022). Integrating self-directed learning in facilitating writers engagement through Grammarly: Exploring the perceptions of premium users. EnJourMe (English Journal of Merdeka). *Culture, Language, and Teaching of English, 7*(1), 32–46. https://doi.org/10.26905/enjourme.v7i1.6849

Williams, J. S., & Lowman, R. L. (2018). The efficacy of executive coaching: An empirical investigation of two approaches using random assignment and a switching-replications design. *Consulting Psychology Journal: Practice and Research, 70*(3), 227–249. https://doi.org/10.1037/cpb0000115

Wilson, J., & Roscoe, R. D. (2019). Automated Writing Evaluation and Feedback: Multiple Metrics of Efficacy. *Journal of Educational Computing Research, 58*(1), 87–125. https://doi.org/10.1177/0735633119830764

Wong, W. L., Muhammad, M. M., Chuah, K. P., Ma'arop, A. H., & Elias, R. (2022). Did you Run the Telegram? Use of Mobile Spelling Checker on Academic Writing. *Multilingual Academic Journal of Education and Social Sciences, 10*(1), 20. http://dx.doi.org/10.46886/MAJESS/v10-i1/7379

Wu, P., Yu, S., & Luo, Y. (2023). The development of teacher feedback literacy in situ: EFL writing teachers' endeavor to human-computer-AWE integral feedback innovation. *Assessing Writing Volume, 57*, 100739. https://doi.org/10.1016/j.asw.2023.100739

Wu, X. (2022). Dynamic evaluation of college English writing ability based on AI technology. Journal of Intelligent Systems, 31(1), 298–309. https://doi.org/10.1515/jisys-2022-0020

Wu, Y. T., & Wang, L. J. (2023). Advancing University EFL Students' Argumentative Essay Writing Performance through Knowledge-Building-based Holistic Instruction. *Educational Technology & Society, 26*(3), 115–128. https://doi.org/10.30191/ETS.202307_26(3).0009

Yamanishi, H., Ono, M., & Hijikata, Y. (2019). Developing a scoring rubric for L2 summary writing: A hybrid approach combining analytic and holistic assessment. *Language Testing in Asia, 9*(13), 1–22. https://doi.org/10.1186/s40468-019-0087-6

Yang, L. (Francoise), Zhang, L. J., & Dixon, H. R. (2023). Understanding the impact of teacher feedback on EFL students' use of self-regulated writing strategies. *Journal of Second Language Writing, 60*. https://doi.org/10.1016/j.jslw.2023.101015

Yuniar, R. F., Widiati, U., & Astuti, U. P. (2019). The Effect of Using Wattpad on Process-Genre Approach towards Writing Achievement in Tertiary Level. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan, 4*(7), 897–905. http://dx.doi.org/10.17977/jptpp.v4i7.12631

Zhai, N., & Ma, X. (2023). The Effectiveness of Automated Writing Evaluation on Writing Quality: A Meta-Analysis. *Journal of Educational Computing Research, 61*(4), 875–900. https://doi.org/10.1177/07356331221127300.

Zhang, J., Ozer, H. Z., & Bayazeed, R. (2020). Grammarly VS face-To-Face at The Writing Centre: ESL Students Writers'Perceptions. *Praxis: A Writing Center Journal, 17*(2), 2–17. https://doi.org/doi: http://dx.doi.org/10.26153/tsw/8523

Zhang, R., & Zou, D. (2022). Types, features, and effectiveness of technologies in collaborative writing for second language learning. *Computer Assisted Language Learning, 35*(9). https://doi.org/10.1080/09588221.2021.1880441

Zheng, X., Luo, L., & Liu, C. (2023). Facilitating Undergraduates' Online Self-Regulated Learning: The Role of Teacher Feedback. *The Asia-Pacific Education Researcher, 32*(6), 805–816. https://doi.org/10.1007/s40299-022-00697-8

Zou, D., Xie, H., & Wang, F. L. (2023). Effects of technology enhanced peer, teacher and self-feedback on students' collaborative writing, critical thinking tendency and engagement in learning. *Journal of Computing in Higher Education, 35*(1), 166–185. https://doi.org/10.1007/s12528-022-09337-y

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 134-149 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20665

Relationship between knowledge, perception of competence and teachers' performance against cyberbullying

Relación entre el conocimiento, la percepción de competencia y la actuación de los profesores frente al ciberacoso

RECEIVED 12/10/2024 ACCEPTED 12/01/2025 PUBLISHED 01/06/2025

Almudena Castellanos Sánchez
ORCID 0000-0002-5498-8603
Universidad Internacional de La Rioja, España
almudena.castellanos@unir.net

Beatriz Ortega-Ruipérez
ORCID 0000-0002-3822-5745
Universidad Rey Juan Carlos, España
beatriz.ortega@urjc.es

Lucía Granados Alós
ORCID 0000-0002-8620-8333
Universidad Internacional de Valencia, España
lucia.granados@professor.universidadviu.com

David Aparisi Sierra
ORCID 0000-0001-5970-2419
Universidad Internacional de La Rioja, España
david.aparisi@unir.net

ABSTRACT

Cyberbullying is a social phenomenon with serious consequences for those who suffer it and the degree of knowledge and action of teachers is key to its prevention and intervention. The objective of this research was to examine the relationship between teachers' knowledge, perception of competence, and performance against cyberbullying in a sample of 258 post-graduate teachers of different nationalities. An ad hoc questionnaire was designed, and its properties were evaluated using the confirmatory factor analysis technique. To evaluate the predictive capacity of the different variables on cyberbullying behaviour, a logistic regression analysis was carried out. The results show that teachers with high scores in knowledge and perceived competence were more likely to act in cases of cyberbullying. These results highlight the need to strengthen teacher training curricula for cyberbullying training to achieve positive interventions against cyberbullying in schools. Future research should continue to examine the predictive factors associated with successful action.

KEYWORDS Cyberbullying; teachers; knowledge; perception of competence; logistic regression.

RESUMEN

El ciberacoso es un fenómeno social con graves consecuencias para quienes lo sufren y el grado de conocimiento y actuación del profesorado es clave para su prevención e intervención. El objetivo de esta investigación fue examinar la relación entre el conocimiento, la percepción de competencia y la actuación de los profesores frente al ciberacoso en una muestra de 258 profesores de postgrado de diferentes nacionalidades. Se diseñó un cuestionario *ad hoc* y se evaluaron sus propiedades mediante la técnica de análisis factorial confirmatorio. Para evaluar la capacidad predictiva de las distintas variables sobre la conducta de ciberacoso, se realizó un análisis de regresión logística. Los resultados muestran que los profesores con puntuaciones altas en conocimientos y competencia percibida eran más propensos a actuar en casos de ciberacoso. Estos resultados ponen de relieve la necesidad de reforzar los planes de estudio de formación del profesorado en materia de ciberacoso para lograr intervenciones positivas contra el ciberacoso en las escuelas. Futuras investigaciones deberían seguir examinando los factores predictivos asociados con el éxito de la acción.

PALABRAS CLAVE Ciberacoso; profesores; conocimiento; percepción de competencia; regresión logística.

1. INTRODUCTION

Cyberbullying is reported as an aggressive, intentional act carried out by a group or individual, using electronic forms of contact, repeatedly and over time against a victim who cannot easily defend him or her-self (Smith et al., 2008). Although the prevalence rate varies from country to country and also depends on the measurement instruments, overall, according to the latest data in the European Union (Cosma et al., 2024), 15% of adolescents (around 1 in 6) have experienced cyberbullying, with the rates closely aligned between boys (15%) and girls (16%). This represents an increase over 2018, from 12% to 15% for boys and 13% to 16% for girls. Although most current studies analyze cyberbullying at the secondary education stage, it has been shown that the roles of cybervictim and cyberbully can arise from the primary education stage, and teacher training and collaboration between families and the school is necessary (Chicote-Beato et al., 2024; Flores et al., 2020).

Of the protective factors identified to decrease the likelihood of cyberbullying include individual factors such as higher self-esteem, strong ties at the family level, a high degree of emotional intelligence, and environmental factors such as residence in safer neighborhoods and a positive school climate including teacher involvement (Kowalski et al., 2019). In this sense, there is a clear gap in the scientific literature when it comes to studying variables that might predict positive teacher involvement in cyberbullying cases. Although studies have been developed that analyse cyberbullying in relation to teachers, most of them focus on studying the perception that teachers have of this phenomenon (Compton et al., 2014; Giménez-Gualdo et al., 2018; Gradinger et al., 2017; Green et al., 2016; Huang & Chou, 2013; Mishna et al., 2020; Monks et al., 2016; Redmond et al., 2018; Samara et al., 2020; Sigal et al., 2012; Yot-Domínguez et al., 2019). To a lesser extent, teacher performance in the face of cyberbullying has been studied (Macaulay et al., 2018; Nappa et al., 2020), teacher knowledge of cyberbullying (Campbell, et al., 2018; Redmond et al., 2020) and even the success of some programmes for teachers to improve coping with cyberbullying has been investigated (Del Rey et al., 2019; Guarini et al., 2019), however, little is known about predictors that might influence teachers' encouragement to take action when they observe or are alerted to cyberbullying at school. In this regard,

some studies that have focused on analysing the victim's perspective show that not all teachers react to cyberbullying (Chaves et al., 2020; Giménez et al., 2018).

It is now known that school climate is related to cyberbullying, a negative climate, a lower sense of belonging to the school by students, increases the likelihood of participating as a bully (Varela et al., 2018; Williams & Guerra, 2007; Wong et al., 2014). Within the school climate, the teacher-student relationship has also been investigated as a protective factor, positive relationships in schools, including school staff in fostering them, help to build a more protective environment for students, however, an adequate level of knowledge about cyberbullying among teachers is necessary for their support to be effective (Olenik-Shemesh et al., 2019; Varela et al., 2018).

Given this lack of empirical research to understand the importance of the teacher's role in cases of cyberbullying, some studies have been emerging that shed light on the factors that influence teachers to intervene in these situations. Sardessai et al. (2021), based on the theory of planned behaviour, formulated by Ajzen and Fishbein (1980), find that attitude, understood as the degree of a person's favourable or unfavourable evaluation of a behaviour, is the variable that best predicts teachers' intention to intervene. Previously, another study suggested that coping with cyberbullying is positively correlated with high levels of teacher empathy (Olenik- Shemesh et al., 2019). However, other variables that could also predict a teacher's intention to intervene in a cyberbullying case, such as subject matter knowledge or self-efficacy, remain unexplained (Sardessai et al., 2021).

It has recently been found that positive family involvement in cyberbullying cases can be predicted from the level of knowledge about cyberbullying, perceived competence or self-efficacy, risk adjustment and attribution of responsibility (Gohal et al., 2023; Martín-Criado et al., 2021). Understanding how people behave when faced with a challenge or problem is one of the frequent topics of study in psychology. In this sense, social cognitive theory (Bandura, 1977) maintains that expectations of personal efficacy determine whether a coping behaviour will be initiated in the face of a problem, understanding self-efficacy as the beliefs that a person has about his or her abilities to perform an action. Numerous studies point to the self-efficacy variable as a significant predictor of behaviour (De Vries et al., 1988; Holloway & Watson, 2002; Schwarzer & Fuchs, 1996). The recent study by Maurya et al. (2023) aimed to study the relationship between parental self-efficacy, communication and cyber victimization and depression in a sample of youth in India. The results showed that young victims of cyberbullying improved their mental health as communication and parental self-efficacy increased. In this sense, the importance of the sense of self-efficacy in dealing with cyberbullying situations is observed. Similarly, considerable research reveals the important role of risk perception in human behaviour (Arezes & Miguel, 2005). The term risk can be defined as the possibility of an undesirable state of reality occurring because of natural events or human activities (Rohrmann & Renn, 2000) and in the social sciences, people's views on risk are often referred to as risk perception. Precisely the fear of that perceived risk predisposes people to act, motivating them to seek protective measures, as maintained by the theory of protective motivation (Rogers, 1975). In turn, responsibility, understood as the demand both to others and to oneself of a response to an interpellation (Crespo & Freire, 2014), has been extensively studied by social psychology. Attributing and assuming responsibility are reciprocally related processes, as can be extracted from Heider's (1958) theory of attribution, all this with the

aim of developing coping strategies. Despite the strategies, which in some cases are implemented, most teachers point out the lack of specific training to intervene (Cerezo & Rubio, 2017), and even to detect cyberbullying even when it affects students in their own classrooms (Montoro & Ballesteros, 2016). In this sense, greater involvement, specific training and intervention of teachers is necessary (Bevilacqua et al., 2017), as well as their training and planning to be able to intervene in the face of cyberbullying (Giménez-Gualdo et al., 2018; Nocentini et al., 2015).

1.1. The present study

Although previous research has shown the relationship between cyberbullying and several variables such as the perception of competence and knowledge in families in cyberbullying situations (Ho et al., 2029; Martín-Criado et al., 2021), there are not many studies that study exactly the relationship between cyberbullying and these variables in teachers. Therefore, the present research pursues two objectives: (1) to study the relationship between the variable's knowledge of cases, perception of competence and teachers' performance in cases of cyberbullying; and (2) to analyse the predictive capacity of knowledge and perception of competence on performance in cases of cyberbullying in a sample of teachers.

Based on the review of previous research, we expect to find differences in the variables studied. Specifically, a statistically significant relationship is expected to be found between the degree of knowledge and action in cases of cyberbullying (hypothesis 1). With regard to perceived competence, a statistically significant and positive relationship is expected to be found between competence and cyberbullying behaviour (hypothesis 2). Finally, teachers' knowledge and perceived competence are expected to be significant predictors of cyberbullying performance (hypothesis 3).

2. MATERIAL AND METHOD

2.1. Participants

A total of 295 teachers from different educational levels participated in the study, of which 37 were eliminated due to errors in their responses to the questionnaire, after the use of a non-probabilistic sampling by convenience sampling. Therefore, the final research sample consisted of 258 active teachers in a graduate course in educational technology and digital competencies, using non-probability sampling by accessibility to select the sample. The average age of the teachers ranged from 29 to 65 years (M = 44.61; SD = 12.25) and the gender distribution was 52% female and 48% male. Using the Chi-square Test of Homogeneity of frequencies distribution, it was proved that there were no statistically significant differences among Sex x Age ($\chi^2 = 3.15$; $p = 0.37$). The country of origin of the teachers was as follows: Spain (20.1%), Ecuador (31.8%) and Colombia (42.6%). All teachers work in different educational stages, from Early Childhood Education (0-6 years) to Higher Education (17 years or more), with the Secondary Education stage (12-16 years) being the most numerous (46.1%), followed by Primary Education (6-12 years), representing 26%. Of the total number of teachers in the sample, 68.2% are working for more than 4 years, 47.7% of the teachers have more than 6 years of experience, but there are also teachers who are working between 1 and 3 years (19.4%) and even some less than one year (12.4%).

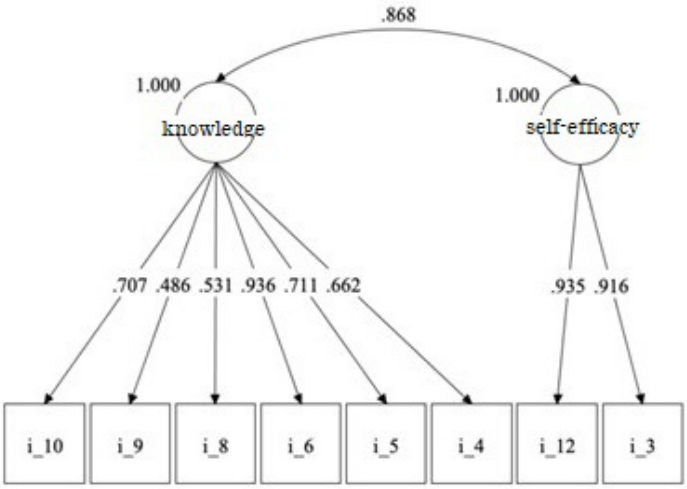
TABLE 1. Distribution of the number of teachers surveyed according to country of residence and specialisation

Teaching specialisation	Country of residence								
	USA	Peru	Spain	Colombia	Ecuador	Belgium	Switzerland	Chile	Dominican Rep.
Students 0-6 years old	-	-	11	3	9	2	-	1	-
Students 6-12 years old	-	1	27	22	17	-	-	-	-
Students 12-16 years old	1	2	7	64	42	-	-	-	3
Students 17 years and older	-	-	7	21	14	-	-	3	0
Total	**1**	**3**	**52**	**110**	**82**	**2**	**1**	**4**	**3**
% Total Participants	**0.4%**	**1.2%**	**20.1%**	**42.6%**	**31.8%**	**0.8%**	**0.4%**	**1.5%**	**1.2%**

2.2. Instruments

The purpose of the research is to analyse the predictive ability of the variables knowledge and perceived competence of teachers on their performance in cases of cyberbullying. Given that cyberbullying is a relatively recent phenomenon, few measurement instruments have been developed that address teachers' perceptions, in this sense, an ad hoc questionnaire is designed based on similar research (Li, 2008) to examine the teachers' perspective on cyberbullying. The questionnaire consists of a total of 8 items relating to the overall construct, although divided into two main dimensions: knowledge about cyberbullying and self-efficacy or competence. The questions referring to the degree of knowledge of cyberbullying were the following: 1: Did you receive training on cyberbullying during your studies? What was involved?, 2: Have you received training on coexistence? What was involved?, 3: Have you received training on cyberbullying at school? What was involved?, 4: Does your school have a Code of Coexistence? What was involved? and 5: Does your school have a protocol against bullying? What was involved?, 6: Does your school have a protocol in case of cyberbullying? What was involved? The two questions referring to perceived competence and self-efficacy were the following: 1: Do you see yourself as needing specific training on cyberbullying? Why?, and 2: Do you know how to deal with cyberbullying in your school? What was involved? The questionnaire has been approved by the Scientific-Ethical Committee of the University with the code PI:035/2021. The questions were formulated in an open-ended manner so as not to condition the answers. These were analysed according to a system of categories developed from the data. The reliability of the category system was determined by the independent scoring of the categories by two judges. The inter-rater reliability of the categories was calculated using the *Kappa* index of inter-rater reliability (Cohen, 1960), applied to each category. Agreement was quite high on all questions, with means ranging from .84 to .99 in the different categories. It should be noted that the questionnaire was validated by a committee of experts prior to its application. Figure 1 shows the different items.

FIGURE 1. Questionnaire items



2.3. Procedure

Once the sample of teachers had been selected, a meeting was held to inform them of the objectives of the study, ask their permission, explain the evaluation instruments, and thank them for their collaboration in answering the questionnaire, which was completed voluntarily and collectively. Anonymity of study participants was ensured by assigning identification numbers to the answer sheets. Researchers were also available during the tests to clarify any doubts and to confirm the correct administration of the questionnaire. Compliance with the test was ensured, taking an average of 50 minutes to complete the test. This research complies with the ethical principles for the conduct of research involving human subjects according to the Declaration of Helsinki of the World Medical Assembly.

2.4. Statistical analysis

Firstly, an exploratory factor analysis was carried out on the 8 items of the questionnaire applied to the sample of teachers, and it was found that the items were grouped into 2 factors: knowledge about cyberbullying and perception of competence and self-efficacy. Once the model was estimated, confirmatory factor analysis was carried out, which made it possible to assess the validity and reliability of each item of the questionnaire. Once it was verified that the estimated fit function reached the minimum, the quality of the model was analysed. To measure the goodness-of-fit, different indices were calculated, which report the extent to which the structure defined by the model parameters reproduces the covariance matrix of the sample data (CFI = .93, RMSEA = .005, GFI = .90).

For the analysis of the relationship between the categorical variable teacher performance and the rest of the variables, Spearman's correlation coefficient was used, since these, in addition to being categorical variables, do not comply with the assumption of normality (Powers & Xie, 2000). The interpretation of Spearman's coefficient ranges from -1 to +1, indicating negative or positive associations respectively, 0 zero, meaning no correlation, but not independence. The interpretation of the Spearman correlation coefficient values follows the following criteria (and considering absolute values): between 0 and .10:

non-existent correlation; between .10 and .29: weak correlation; between .30 and .50: moderate correlation; and between .50 and 1: strong correlation.

Secondly, a logistic regression analysis was carried out using the forward stepwise procedure following the Wald statistic to assess the predictive capacity of the study variables on acting in the face of cyberbullying. The variable "action against cyberbullying" (yes/no) by teachers was dichotomized based on the answers given to the questions. In this case, the probability of the event occurring (acting against cyberbullying) was analysed using the Odds Ratio (OR), which is interpreted as follows: if the OR is greater than one, for example 2, for every time the event occurs in the presence of the independent variable, it will occur two times if this variable is present. Conversely, if the OR is less than one, e.g. 0.5, the probability of the event occurring in the absence of the independent variable will be greater than in its presence (Aparisi et al., 2021). Nagelkerke's R^2 was used to assess the quality of the proposed models and their fit, indicating the percentage of variance explained by the model (Nagelkerke, 1991) and the percentage of cases correctly classified by the model or the predicted effectiveness.

Finally, for the static analysis, we used the program SPSS version 26.0.

3. RESULTS

3.1. Differences in teachers' knowledge, perceived competence, and performance against cyber-bullying cases

As can be seen in Table 2, there is a positive and statistically significant relationship between the variable teacher performance in cyberbullying cases and knowledge of cyberbullying cases ($p < .05$), with the intensity of the association between the variables being weak. In addition, there is also a positive and statistically significant relationship between the variable's teacher performance and perceived competence ($p < .05$), with a weak association between the variables. Finally, there is also a positive and statistically significant relationship between the variable's teacher knowledge of the cases and perception of competence, the intensity of the relationship between the variables being moderate.

TABLE 2. Results derived from Spearman's correlation between the performance variables

			Performance	Knowledge	Perception of competence
Spearman's Rho	Performance	Correlation coefficient	1	.21**	.14*
		Sig. (bilateral)	.	.00	.02
		N	258	258	258
	Knowledge	Correlation coefficient	.21**	1	.36**
		Sig. (bilateral)	.00	.	.00
		N	258	258	258
	Perception of competence	Correlation coefficient	.14*	.36**	1
		Sig. (bilateral)	.02	.00	.
		N	258	258	258

Note. **. Correlation is significant at the .01 level. *. Correlation is significant at the .05 level.

3.2. Predictive ability of teachers' knowledge and perceived competence in cyberbullying behaviour

Two predictive models were obtained from the results of the logistic regression analysis from teachers' performance in cyberbullying cases based on knowledge and perception of their competence (Table 3), each of them correctly classifying 79.4% of the cases ($\chi^2 = 20.93$; p = .00) and 78.2% ($\chi^2 = 12.21$; p = .00), respectively. The fit value (R^2 Nagelkerke) of both models was .12 and .07, respectively. Odds Ratio (OR) indicate that teachers are 57% more likely to act on cyberbullying cases as their score on the knowledge scale increases by one unit and 3.0 times more likely to act on cyberbullying cases as their score on the perceived competence variable increases by one unit.

TABLE 3. Results derived from binary logistic regression for the probability of acting on cyberbullying cases

Model	Predictor variable	B	E.T.	Wald	*p*	*OR*	I.C. 95%
Performance	Knowledge	. 45	.10	19.15	.00	1.57	1.28-1.93
	Constant	-2.45	.33	53.26	.00	.08	
	Perception of competence	1.09	.31	11.92	.00	3.00	1.60-5.60
	Constant	-2.60	.43	36.06	.00	.07	

Note. B = coefficient; S.E. = standard error; *p* = probability; *OR* = odds ratio; C.I. = 95% confidence interval

4. DISCUSSION

The aim of the present study was twofold. On the one hand, to study the relationship between the variable's knowledge of cases, perception of competence and teachers' performance in cases of cyberbullying; and on the other hand, to analyse the predictive capacity of knowledge and perception of competence on performance in cases of cyberbullying in a sample of teachers.

Specifically, a statistically significant relationship is expected to be found between the degree of knowledge and action in cases of cyberbullying (hypothesis 1). With regard to perceived competence, a statistically significant and positive relationship is expected to be found between competence and cyberbullying behaviour (hypothesis 2). Finally, teachers' knowledge and perceived competence are expected to be significant predictors of cyberbullying performance (hypothesis 3).

The results obtained confirmed the three hypotheses of the study, as a positive and statistically significant relationship was obtained between the variables knowledge of cyberbullying cases, perception of competence and teachers' performance. Furthermore, teachers' knowledge of cyberbullying cases and their perceived competence were statistically significant predictors of teachers' performance in cyberbullying cases. In line with previous studies, the results suggest that perceived competence and knowledge are key aspects of teachers' performance in dealing with cyberbullying situations. In this regard, the study by Sardessai-Nadkarni et al. (2021) conducted with a sample of 402 teachers in India found that attitude, subjective norms, and perceived control over the situation explained 40.9% of the variance in intentions. Another study with 644 teachers in Israel showed that coping with cyberbullying is positively correlated with high levels of

empathy, communication with students and self-efficacy (Olenik-Shemesh et al., 2019). The aim of the research of Hurtubise (2021) was to explore how teachers' perceptions affected their likelihood of responding to varied cyberbullying scenarios (e.g., whether at home or school). Using multilevel modelling, this study investigated the relationships between teachers' likelihood of response and key psychological factors and background characteristics, drawing on a convenience sample of 212 new and experienced teachers from England and the United States. Some of these factors include valence (severity of cyberbullying), expectancy (level of teacher self-confidence), and instrumentality (confidence in selected task). Findings show that valence, expectancy, and location of the cyberbullying were statistically significant predictors of teachers' likelihood of response to situations of cyberbullying.

Regarding teachers' knowledge about cyberbullying, other research warns, for example, of the importance of educating teachers about the need for cyberbullying prevention programmes (Hirschstein et al., 2007; Stauffer et al., 2012), or about school climate as a preventive variable (Cohen et al., 2009). In this sense, the study by Sidera et al. (2019), with a sample of 220 Pre-school and Primary School teachers, showed that 24.1% felt trained to deal with a situation of cyberbullying or traditional bullying, while the vast majority, 61.7%, were not sure they had the necessary skills and 14.2% admitted not feeling qualified to deal with a case of these characteristics. Consequently, the authors point out the importance of training teachers in the field of cyberbullying in terms of protocols and prevention programmes adapted to the reality of each school. A meta-analysis study of nineteen programs reviewed from 2700 selected articles on cyberaggression and cybervictimization showed that only programs involving interpersonal interactions and stakeholder (including teachers) action demonstrated superior program effectiveness (Lan et al., 2022). Therefore, teacher training should be based on educational programs that allow teachers to improve their understanding of the phenomenon of cyberbullying and be able to participate in the analysis of the causes, the extent of the consequences and the design of educational actions to address it (Barlett, 2017) and combine the implementation of strategies designed by experts with personalized actions according to the characteristics of the students and the school's own culture (Del Rey et al., 2019). Recent research points to the importance of training teachers in social-emotional skills in the prevention of cyberbullying that impact the sense of self-efficacy when intervening when the problem manifests itself (Llorent & Núñez-Flores, 2024; Schoeps et al., 2018). Teachers with good social and emotional competencies will be able to develop the same competencies more effectively in their students, thereby preventing and reducing youth involvement in cyberbullying. In this sense, the study by Gabarda et al. (2022) concluded, in a sample of 653 university professors, the need for training in teacher competencies as well as the importance of the perception of the problem and the sense of self-efficacy in dealing with cyberbullying situations among students.

According to the third hypothesis of the study, perceived competence and self-efficacy is a significant predictor of the likelihood of cyberbullying intervention (Williford & Depaolis, 2016). According to Klassen and Tze (2014), self-efficacy, in the case of teachers, increases persistence in working with difficult students, influences teaching practices, as well as enthusiasm and commitment. The same is true for students: those who are confident in their competence in cyberbullying situations regulate their behaviour more successfully (Bussey et al., 2020). Teacher apathy, understood as the lack of interest or poor quality of the teacher's performance, in relation to the phenomena of indiscipline, influences cyberbullying, as the research carried out by

Ortega et al. (2013) also concludes. Many studies demonstrate that educators have knowledge of cyberbullying and perceive it as an issue in their schools, yet the majority of them lack the necessary skills and training to deal with it (Çırak & Demirkan, 2023; Eden et al., 2013; Fredrick et al., 2023; Tomczyk & Włoch, 2019).

5. CONCLUSIONS

In conclusion, this study provides relevant information on the phenomenon of cyberbullying and teacher involvement. In this sense, this research confirms the existence of statistically significant differences in teachers' knowledge, perception of competence and performance in the face of cyberbullying.

In addition, teachers' knowledge of the cases and perceived competence were statistically significant predictors of cyberbullying behaviour, as teachers with high scores on knowledge and perceived competence were more likely to act on cyberbullying. Developing resilience to cyberbullying involves conducting educational awareness programs (Ng et al., 2022), exercising self-efficacy, and utilizing technological knowledge (Achuthan et al., 2023). Acting effectively against cyberbullying in educational environments requires more digitally competent teachers in knowledge, skills (social and technical) and attitudes, especially as related to digital safety (Torres-Hernández & Gallego-Arrufat (2022). Teachers have a very important role in the education of students, not only at the academic level but also at the personal and emotional level, providing support. Teachers have the responsibility to show strong leadership within the educational system, to improve coexistence and to attend to all issues that occur in the school environment (Epstein & Kazmierczak, 2006). Likewise, educational institutions have the responsibility to prepare future professionals to be more competent in dealing with cyberbullying (Musset, 2010). Continuing education and training of future teachers in the university context will provide a valuable platform to promote school culture and attitudes in hopes of reducing cyberbullying situations.

5.1. Limitations and future lines of research

This research is not without some limitations. Firstly, although the technique used to select the sample guarantees its representativeness, it would be necessary to confirm the results found in this study at other educational levels (for example, in higher or university education). In this sense, it would be interesting to consider future longitudinal studies to verify the relationship between the variables in the long term. It is also worth bearing in mind the possible biases or limitations of teachers' self-reported data, taking into account the effects of social desirability bias. In this sense, an in-depth qualitative evaluation of the responses would make it possible to control for such limitations. Secondly, it is important to consider the mediating effect of other personal variables (such as self-esteem, emotional intelligence), as well as social and family variables that may influence the relationship between cyberbullying and the variables under study. It would also be interesting to know the differences found in the present study according to the sex of the teachers. In this sense, it has been observed that women have higher scores in empathy, related to the feeling of self-efficacy in the classroom (Goroshit & Hen, 2016). Finally, this research aims to find out the predictive capacity of knowledge and the perception of competence on the performance of teachers in cases of cyberbullying. While it is reasonable to assume that there is an influential relationship between the variables,

future research could test this relationship by developing two structural equation models to test the more stable hypothesis or, if so, what is the strength of the relationship in both models.

From a practical point of view, the results of this study highlight the need to design and implement educational intervention programmes aimed at improving teachers' perceptions of competence. In this sense, self-efficacy beliefs are formed from four sources of information postulated by Bandura (1977): mastery experiences, verbal persuasion, vicarious experience, and affective states. The recent study by Gümüş et al. (2023) concluded that awareness of digital data security predicts sensitivity to cyberbullying.

Secondly, research has shown that teachers who are aware of cyberbullying situations are the ones who can finally do something to prevent or solve these cases. Therefore, it is essential to improve the channels of information and reporting of possible situations of cyberbullying among students. As stated by Romera et al. (2016), teachers and counsellors need clarifying knowledge and action models to manage groupings, work on improving classroom climate, developing social activities, analysing classroom relationships, and establishing interpersonal links. Finally, the collaboration of families and other educational professionals is essential to end this problem and reduce the negative effects on victims (Hellfeldt et al., 2020). In this sense, school and family are the two most important socialisation agents in students' lives and should cooperate closely by educating empathy (Zhang et al., 2020).

Teachers stated that knowledge about cyberbullying cases and training is essential to increase online self-efficacy. According to previous studies, teachers believe that they lack the necessary resources and that their professional training has not adequately equipped them to handle cyberbullying concerns (Fredrick et al., 2023). This finding can be very useful in creating and improving teacher training in Higher Education. Teachers with higher levels of digital competence are more adept at detecting and acting on cyberbullying cases. To successfully implement digital technologies in the classroom and enhance the quality of teaching, both educational institutions and policymakers need to establish appropriate regulations that encourage teachers to improve their digital skills, actively engage in digital technology education, enhance their perspectives and self-efficacy in its use, and foster the growth of digital competencies (Macaulay et al., 2018). According to some studies, although teachers claim that being trained in cyberbullying knowledge would be effective (Pelfrey & Weber, 2014), they lack confidence in their ability to identify and manage the problem (Barnes et al., 2012). Therefore, both training in improving the detection of possible cases of cyberbullying and improving the sense of self-efficacy through teacher training are two fundamental tools in solving the problem of cyberbullying.

6. REFERENCES

Achuthan, K., Nair, V. K., Kowalski, R., Ramanathan, S., & Raman, R. (2022). Cyberbullying research—Alignment to sustainable development and impact of COVID-19: Bibliometrics and science mapping analysis. *Computers in Human Behavior*, *140*. https://doi.org/10.1016/j.chb.2022.107566

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.

Aparisi, D., Delgado, B., Bo, R. M., & Martínez-Monteagudo, M. C. (2021). Relationship between Cyberbullying, Motivation and Learning Strategies, Academic Performance, and the

Ability to Adapt to University. *International Journal of Environmental Research and Public Health, 18,* 10646. https://doi.org/10.3390/ijerph182010646

Arezes, P. M., & Miguel, A. S. (2005). Hearing protection use in industry: The role of risk perception. *Safety Science, 43*(4), 253-267. https://doi.org/10.1016/j.ssci.2005.07.002

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295x.84.2.191

Barlett, C. P. (2017). From theory to practice: Cyberbullying theory and its application to intervention. *Computers in Human Behavior*, *72*, 269-275. https://doi.org/10.1016/j.chb.2017.02.060

Barnes, A., Cross, D., Lester, L., Hearn, L., Epstein, M., & Monks, H. (2012). The invisibility of covert bullying among students: Challenges for school intervention. *Journal of Psychologists and Counsellors in Schools*, *22*(2), 206–226. https://doi.org/10.1017/jgc.2012.27

Bevilacqua, L., Shackleton, N., Hale, D., Allen, E., Bond, L., Christie, D., & Viner, R.M. (2017). The role of family and school-level factors in bullying and cyberbullying: a cross-sectional study. *BMC Pediatrics*, *17*, 160 https://doi.org/10.1186/s12887-017-0907-8

Bussey, K., Luo, A., Fitzpatrick, S., & Allison, K. (2020). Defending victims of cyberbullying: The role of self-efficacy and moral disengagement. *Journal of School Psychology*, *78*, 1–12. https://doi.org/10.1016/j.jsp.2019.11.006

Campbell, M., Whiteford, C., & Hooijer, J. (2018). Teachers' and parents' understanding of traditional and cyberbullying. *Journal of School Violence, 18*(3), 388–402. https://doi.org/10.1080/15388220.2018.1507826

Cerezo, F., y Rubio, F.J. (2017). Medidas relativas al acoso escolar y ciberacoso en la normativa autonómica española. Un estudio comparativo. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, *20*(1), 113-126. http://dx.doi.org/10.6018/reifop/20.1.253391

Chaves, A. L., Morales, M. E., & Villalobos, M. (2020). Cyberbullying from the students' perspective: "What we live, see and do". *Revista Electrónica Educare*, *24* (1), 1-29. https://doi.org/10.15359/ree.24-1.3

Chicote-Beato, M., González-Víllora, S., Bodoque-Osma, A., & Navarro, R. (2024). Cyberbullying intervention and prevention programmes in Primary Education (6 to 12 years): A systematic review. *Aggression and Violent Behavior*, *77*, 101938. https://doi.org/10.1016/j.avb.2024.101938

Çırak, Y., & Demirkan, Ö. (2023). Comparison of Teachers' Online Technologies Self-Efficacy and Cyberbullying Awareness (Cyprus Sample). *International Journal of Innovative Approaches in Education*, *7*, 166-183. https://doi.org/10.29329/ijiape.2023.627.1

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46. https://doi.org/10.1177/001316446002000104

Cohen, J., McCabe, L., Michelli, N. M., & Pickeral, T. (2009). School climate: Research, policy, practice, and teacher education. *Teachers College Record*, *111*(1), 180–213. https://doi.org/10.1177/016146810911100108

Compton, L., Campbell, M.A., & Mergler, A. (2014). Teacher, parent and student perceptions of the motives of cyberbullies. *Social Psychology of Education, 17,* 383–400. https://doi.org/10.1007/s11218-014-9254-x

Cosma, A., Molcho, M., & Pickett, W. (2024). *A focus on adolescent peer violence and bullying in Europe, central Asia and Canada. Health Behaviour in School-aged Children international report from the 2021/2022 survey* (Vol.2). World Health Organization. https://iris.who.int/handle/10665/376323

Crespo, E., & Freire, J.C. (2014). The attribution of responsibility: from cognition to the subject. *Psicología y Sociedade*, *26*(2), 271-279. https://doi.org/10.1590/S0102-71822014000200004

De Vries, H., Dijkstra, M., & Kuhlman, P. (1988). Self-efficacy: the third factor besides attitude and subjective norm as a predictor of behavioural intentions. *Health Education Research*, *3*(3), 273-282. https://psycnet.apa.org/doi/10.1093/her/3.3.273

Del Rey, R. Ortega-Ruiz, R., y Casas, J. A. (2019). Asegúrate: An intervention program against cyberbullying based on teachers' commitment and on design of its instructional materials. *International Journal of Environmental Research and Public Health*, *16*(3), 434. https://doi.org/10.3390/ijerph16030434

Eden, S., Heiman, T., & Olenik-Shemesh, D. (2013). Teachers' perceptions, beliefs and concerns about cyberbullying. *British Journal of Educational Technology*, *44*(6), 1036-1052. https://doi.org/10.1111/j.1467-8535.2012.01363.x

Epstein, A., & Kazmierczak, J. (2006). Cyber bullying: What teachers, social workers, and administrators should know. *Illinois Child Welfare*, *3*(1–2), 41–51.

Flores, R., Miedes, A., y Oliver, M. (2020). Efecto de un programa de prevención de ciberacoso integrado en el currículum escolar de Educación Primaria. *Revista de Psicodidáctica*, *25*(1), 23-29. https://doi.org/10.1016/j.psicod.2019.08.001

Fredrick, S. S., Coyle, S., & King, J. A. (2023). Middle and high school teachers' perceptions of cyberbullying prevention and digital citizenship. *Psychology in the Schools*, *60*(6), 1958-1978. https://doi.org/10.1002/pits.22844

Gabarda, C., Cuevas, N., Martí, A., López-Agustí, A. I., y Rodríguez-Martí, A. (2022). Percepción del profesorado ante el Cyberbullying y detección de necesidades formativas del perfil competencial docente. *REIDOCREA*, *11*(38), 458-466. http://dx.doi.org/10.30827/Digibug.76395

Giménez, A.M, Arnaiz, P., Cerezo, F., & Prodócimo, E. (2018). Teachers' and students' perceptions of cyberbullying. Intervention strategies in primary and secondary school. *Comunicar*, *56*(26), 29-38. https://doi.org/10.3916/C56-2018-03

Giménez-Gualdo, A., Arnaiz-Sánchez, P., Cerezo-Ramírez, F., y Prodócimo, E. (2018). Percepción de docentes y estudiantes sobre el ciberacoso. Estrategias de intervención y afrontamiento en Educación Primaria y Secundaria. *Comunicar*, *56*, 29-38. https://doi.org/10.3916/C56-2018-03

Gohal, G., Alqassim, A., Eltyeb, E., Rayyani, A., Hakami, B., Al Faqih, A., Hakami, A., Qadri, A., & Mahfouz, M. (2023). Prevalence and related risks of cyberbullying and its effects on adolescent. *BMC psychiatry*, *23*(1), 39. https://doi.org/10.1186/s12888-023-04542-0

Goroshit, M., & Hen, M. (2016). Teachers' empathy: can it be predicted by self-efficacy? *Teachers and Teaching*, *22*(7), 805–818. https://doi.org/10.1080/13540602.2016.1185818

Gradinger, P., Strohmeier, D. & Spiel, C. (2017). Parents' and Teachers' Opinions on Bullying and Cyberbullying Prevention: The Relevance of Their Own Children's or Students' Involvement. *Zeitschrift für Psychologie, 225*(1), 76-84. https://doi.org/10.1027/2151-2604/a000278

Green, V. A., Johnston, M., Mattioni, L., Prior, T., Harcourt, S., & Lynch, T. (2016). Who is responsible for addressing cyberbullying? Perspectives from teachers and senior managers. *International Journal of School & Educational Psychology, 5*(2), 100–114. https://doi.org/10.1080/21683603.2016.1194240

Guarini, A., Menin, D., Menabò L., & Brighi A. (2019). RPC Teacher-Based Program for Improving Coping Strategies to Deal with Cyberbullying. *International Journal of Environmental Research and Public Health, 16*(6), 948. https://doi.org/10.3390/ijerph16060948

Gümüş, M. M., Çakır, R., & Korkmaz, Ö. (2023). Investigation of pre-service teachers' sensitivity to cyberbullying, perceptions of digital ethics and awareness of digital data security. *Education and Information Technologies*, 28, 14399-14421. https://doi.org/10.1007/s10639-023-11785-7

Heider, F. (1958). *The Psychology of Interpersonal Relations*. Wiley.

Hellfeldt, K., López-Romero, L., & Andershed, H. (2020). Cyberbullying and Psychological Well-being in Young Adolescence: The Potential Protective Mediation Efects of Social Support from Family, Friends, and Teachers. *International of Environmental Research and Public Health*, 17, 45. https://doi.org/10.3390/ijerph17010045

Hirschstein, M. K., Edstrom, L.V.S., Frey, K. S., Snell, J. L., & MacKenzie, E. P. (2007). Walking the talk in bullying prevention: Teacher implementation variables related to initial impact of the Steps to Respect program. *School Psychology Review*, *36*(1), 3-21. https://doi.org/10.1080/02796015.2007.12087949

Ho, S. S., Lwin, M. O., Yee, A. Z. H., Sng, J. R. H., & Chen, L. (2019). Parents' responses to cyberbullying effects: how third-person perception influences support for legislation and parental mediation strategies. *Computers in Human Behavior*, *92*, 373-380. https://doi.org/10.1016/j.chb.2018.11.021

Holloway, A., & Watson, H. E. (2002). Role of self-efficacy and behaviour change. *International Journal of Nursing Practice*, *8*, 2, 106-115. https://doi.org/10.1046/j.1440-172x.2002.00352.x

Huang, Y.-Y., & Chou, C. (2013). Revisiting cyberbullying: Perspectives from Taiwanese teachers. *Computers & Education, 63,* 227–239. https://doi.org/10.1016/j.compedu.2012.11.023

Hurtubise, P. (2021). Teachers' perceptions of cyberbullying: a comparative multilevel modelling approach. *Cambridge Educational Research e-Journal*, *8*, 52-62. https://doi.org/10.17863/CAM.76201

Klassen, R. M., & Tze, V. M. C. (2014). *Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. Educational Research Review, 12,* 59-76. https://doi.org/10.1016/j.edurev.2014.06.001

Kowalski, R. M., Limber, S. P., & McCord, A. (2019). A developmental approach to cyberbullying: Prevalence and protective factors. *Aggress Violent Behav*, 45, 20-32. https://doi.org/10.1016/j.avb.2018.02.009

Lan, M., Law, N., & Pan, Q. (2022). Effectiveness of anti-cyberbullying educational programs: A socio-ecologically grounded systematic review and meta-analysis. *Computers in Human Behavior*, *130*, 107200. https://doi.org/10.1016/j.chb.2022.107200

Li, Qing. (2008). Cyberbullying in schools: An examination of preservice teachers' perception. *Canadian Journal of Learning and Technology*, *34*. https://doi.org/10.21432/T2DK5G

Llorent, V. J., & Núñez-Flores, M. (2024). Social and emotional competencies of teachers to prevent and reduce cyberbullying. In A. L. C. Fung (Ed.), *Cyber and Face-to-Face Aggression and Bullying among Children and Adolescents* (pp. 1-14). Routledge.

Macaulay, P., Betts, L., Stiller, J., & Kellezi, B. (2018). Perceptions and responses towards cyberbullying: A systematic review of teachers in the education system. *Aggression and Violent Behavior*, *43,* 1-12. https://doi.org/10.1016/j.avb.2018.08.004

Martín-Criado, J. M., Casas, J. A., & Ortega-Ruiz, R. (2021). Parental Supervision: Predictive Variables of Positive Involvement in Cyberbullying Prevention. *International Journal of Environmental Research and Public Health, 18*(4), 1562. https://doi.org/10.3390/ijerph18041562

Maurya, C., Muhammad, T., Das, A. Fathah, A., & Dhillon, P. (2023). The role of self-efficacy and parental communication in the association between cyber victimization and depression among adolescents and young adults: a structural equation model. *BMC Psychiatry*, *23*, 337. https://doi.org/10.1186/s12888-023-04841-6

Mishna, F., Sanders, J. E., McNeil, S., Fearing, G., & Kalenteridis, K. (2020). "If Somebody is Different": A Critical Analysis of Parents', Teachers' and Students' Perspectives on Bullying and Cyberbullying. *Children and Youth Services Review, 118,* 105366. https://doi.org/10.1016/j.childyouth.2020.105366

Monks, C. P., Mahdavi, J., & Rix, K. (2016). The emergence of cyberbullying in childhood: Parent and teacher perspectives. *Psicología Educativa, 22*(1), 39-48. https://doi.org/10.1016/j.pse.2016.02.002

Montoro, E., y Ballesteros, M. (2016). Competencias docentes para la prevención del ciberacoso y delito de odio en Secundaria. *Relatec*, *15*(1), 131-143. https://doi.org/10.17398/1695-288X.15.1.131

Musset, P. (2010). Initial teacher education and continuing training policies in a comparative perspective. *OECD education working papers*, *48*, 1-50. *OECD.*

Nagelkerke, N. J. D. (1991). A note on a general definition of the coefficient of determination. *Biometrika, 78*(3), 691–692. https://doi.org/10.1093/biomet/78.3.691

Nappa, M. R., Palladino, B. E., Nocentini, A., & Menesini, E. (2020). Do the face-to-face actions of adults have an online impact? The effects of parent and teacher responses on cyberbullying among students. *European Journal of Developmental Psychology, 18*(6), 798–813. https://doi.org/10.1080/17405629.2020.1860746

Ng, E. D., Chua, J. Y. X., & Shorey, S. (2022). The Effectiveness of Educational Interventions on Traditional Bullying and Cyberbullying Among Adolescents: A Systematic Review and Meta-Analysis. *Trauma, Violence, & Abuse*, *23*(1), 132–151. https://doi.org/10.1177/1524838020933867

Nocentini, A., Zambuto, V., & Menesini, E. (2015). Anti-bullying programs and information and Communication Technologies (ICT): A systematic review. *Aggression and Violent Behavior*, *23*, 52-60. https://doi.org/10.1016/j.avb.2015.05.012

Olenik-Shemesh, D., Heiman, T., y Keshet, N. S. (2019). Factors that affect teachers' coping with cyberbullying: implications for teacher education programs. *Educación Creativa*, *10*, 3357-3371. https://doi.org/10.4236/ce.2019.1013258

Ortega, R., Rey, R. D., y Casas, J. A. (2013). School coexistence: key in the prediction of bullying. *Revista Iberoamericana de Evaluación Educativa*, *6*(2), 91-102. https://doi.org/10.15366/riee2013.6.2.004

Pelfrey, W., & Weber, N. (2014). Student and School Staff Strategies to Combat Cyberbullying in an Urban Student Population. *Preventing School Failure: Alternative Education for Children and Youth*, *59*, 1-10. https://doi.org/10.1080/1045988X.2014.924087

Powers, D. A., & Xie, Y. (2000). *Statistical Method for Categorical Data Analysis*. Academic Press.

Redmond, P., Lock, J. V., & Smart, V. (2018). Pre-service teachers' perspectives of cyberbullying. *Computers & Education, 119*, 1-13. https://doi.org/10.1016/j.compedu.2017.12.004

Redmond, P., Lock, J. V., & Smart, V. (2020). Developing a cyberbullying conceptual framework for educators. *Technology in Society, 60,* 101223. https://doi.org/10.1016/j.techsoc.2019.101223

Rogers, R. W. (1975). A Protection Motivation Theory of Fear Appeals and Attitude Change1. *The Journal of Psychology, 91*(1), 93–114. https://doi.org/10.1080/00223980.1975.9915803

Rohrmann, B., & Renn, O. (2000). Risk perception research. En O. Renn, B. & Rohrmann (Eds.). *Cross-cultural risk perception,* (pp. 11-53). Springer. https://doi.org/10.1007/978-1-4757-4891-8_1

Romera, E. M., Cano, J. J., García, C. M., y Ortega, R. (2016). Cyberbullying: Social Competence, Motivation and Peer Relationships. *Comunicar*, *48*(24), 71-79. https://doi.org/10.3916/C48-2016-07

Samara, M., Da Silva Nascimento, B., El Asam, A., Smith, P., Hammuda, S., Morsi, H. & Al-Muhannadi, H. (2020). Practitioners' perceptions, attitudes, and challenges around bullying and cyberbullying. *International Journal of Emotional Education, 12*(2), 8-25.

Sardessai, A.A., Mclaughlin, B., & Sarge, M.A. (2021). Examining teachers' intentions to intervene: Formative research for school-based cyberbullying interventions in India. *Psychology in the Schools*, *58* (12), 2328-2344. https://doi.org/10.1002/pits.22595

Schoeps, K., Villanueva, L., Prado-Gascó, V. J., & Montoya-Castilla, I. (2018). Development of Emotional Skills in Adolescents to Prevent Cyberbullying and Improve Subjective Well-Being. *Frontiers in Psychology*, *9*, 2050. https://doi.org/10.3389/fpsyg.2018.02050

Schwarzer, R., & Fuchs, R. (1996). Self-efficacy and health behaviors. In M. Conner & P. Norman (Eds.), *Predicting health behavior: Research and practice with social cognition models* (pp. 163–196). Open University Press.

Sidera, F., Rostan, C., Serrat, E., y Ortiz, R. (2019). Teachers facing situations of bullying and cyberbullying at school. *International Journal of Developmental and Educational Psychology INFAD, Revista de Psicología, 3* (1), 421-434. https://doi.org/10.17060/ijodaep.2019.n1.v3.1515

Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry*, *49*, 376–385. https://doi.org/10.1111/j.1469-7610.2007.01846.x

Stauffer, S., Heath, M. A., Coyne, S. M., & Ferrin, S. (2012). High school teachers' perceptions of cyberbullying prevention and intervention strategies. *Psychology in the Schools*, *49*(4), 352–367. https://doi.org/10.1002/pits.21603

Tomczyk, L., & Włoch, A. (2019). Cyberbullying in the light of challenges of school-based prevention. International Journal of Cognitive Research in Science, Engineering and Education, 7(3), 13-26. https://doi.org/10.5937/IJCRSEE1903013T

Torres-Hernández, N., & Gallego-Arrufat, M. J. (2022). Indicators to assess preservice teachers' digital competence in security: A systematic review. *Education and Information Technologies*, 27, 8583–8602. https://doi.org/10.1007/s10639-022-10978-w

Varela, J. J., Zimmerman, M. A., Ryan, A. M., & Stoddard, S.A. (2018). Cyberbullying Among Chilean Students and the Protective Effects of Positive School Communities. *Journal of School Violence*, *17*(4), 430-442. https://doi.org/10.1080/15388220.2017.1358640

Williams, K. R., & Guerra, N. G. (2007). Prevalence and Predictors of Internet Bullying. *Journal of Adolescent Health, 41*(6), 14-21. https://doi.org/10.1016/j.jadohealth.2007.08.018

Williford, A., & Depaolis, K. J. (2016). Predictors of cyberbullying intervention among elementary school staff. *Psychology in the Schools, 53*(10), 1032–1044. https://psycnet.apa.org/doi/10.1002/pits.21973

Wong, D. S. W., Chan, H. C., & Cheng, C. H. K. (2014). Cyberbullying perpetration and victimization among adolescents in Hong Kong. *Children and Youth Services Review*, *36*, 133–140. https://doi.org/10.1016/j.childyouth.2013.11.006

Yot-Domínguez, C., Guzmán Franco, M. D., & Duarte Hueros, A. (2019). Trainee Teachers' Perceptions on Cyberbullying in Educational Contexts. *Social Sciences, 8*(1), 21. https://doi.org/10.3390/socsci8010021

Zhang, X., Han, Z., & Ba, Z. (2020). Cyberbullying Involvement and Psychological Distress among Chinese Adolescents: The Moderating Effects of Family Cohesion and School Cohesion. *International Journal of Environmental Research and Public Health*, *17*(23), 8938. https://doi.org/10.3390/ijerph17238938

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 150-168 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20559

The effect of gamified flipped learning on Malaysian fifth-grade students' academic achievement and learning experience in science

El efecto del aprendizaje invertido gamificado en el rendimiento académico y la experiencia de aprendizaje en ciencias de los estudiantes de quinto grado en Malasia

RECEIVED 26/09/2024 ACCEPTED 10/01/2025 PUBLISHED 01/06/2025

Noor Lydia Russali
ORCID 0009-0001-1424-0828
Universiti Teknologi Malaysia, Malaysia
lydiarussali@gmail.com

Zaidatun Tasir
ORCID 0000-0002-6340-3099
Universiti Teknologi Malaysia, Malaysia
p-zaida@utm.my

ABSTRACT

This study investigated the effect of gamified flipped learning on the academic achievements and learning experiences of fifth-grade students. A quasi-experimental design was employed, with a control group receiving conventional instruction and a treatment group engaging in gamified flipped learning. Forty fifth-grade students were randomly selected, with 20 students in each group. Data were collected through pre- and post-tests, along with interviews, and analysed using descriptive statistics and the Quade-ANCOVA non-parametric test. The results revealed a significant improvement in the academic achievement of the gamified flipped learning group ($p < 0.05$) with a large effect size (partial eta squared = .626) compared to the control group. Interview data further indicated that students in the treatment group experienced enhanced learning, reporting increased excitement, focus, and ease of understanding through gamified activities. These findings underscore the potential of gamified flipped learning to not only improve academic performance but also create engaging and effective learning environments tailored to the needs of Generation Z students. This study contributes valuable insights into the role of gamified flipped learning in primary science education, emphasizing its capacity to transform traditional pedagogies and better support student learning outcomes.

KEYWORDS Gamification; flipped learning; gamified flipped learning; primary school education; academic achievement.

RESUMEN

Este estudio investigó el efecto del aprendizaje invertido gamificado en los logros de los estudiantes de quinto grado. Se utilizó un diseño cuasi-experimental, con un grupo de control que recibió el aprendizaje convencional y un grupo experimental que participó en el aprendizaje invertido gamificado. Se seleccionaron aleatoriamente cuarenta estudiantes de quinto grado, con 20 estudiantes en cada grupo. Los datos se recopilaron mediante pruebas previas y posteriores junto

con entrevistas, y se analizaron utilizando estadísticas descriptivas y la prueba no paramétrica Quade-ANCOVA. El análisis reveló una mejora significativa en los logros del grupo de aprendizaje invertido gamificado ($p < 0.05$) con un tamaño de efecto grande (eta cuadrado parcial = .626) en comparación con el grupo de control. Los datos de las entrevistas indicaron que los estudiantes en el grupo experimental tuvieron experiencias de aprendizaje positivas, informando un aumento en la emoción, el enfoque y la facilidad de aprendizaje a través de actividades gamificadas. Estos hallazgos sugieren que integrar la gamificación con el aprendizaje invertido puede mejorar el logro y el compromiso de los estudiantes, y se recomienda que los educadores continúen incorporando elementos gamificados en el aprendizaje invertido para satisfacer mejor las necesidades de aprendizaje de los estudiantes de la Generación Z.

PALABRAS CLAVE Gamificación; aprendizaje invertido; aprendizaje invertido gamificado; escuela primaria; experiencia de aprendizaje.

1. INTRODUCTION

Primary school science education plays a critical role in shaping students' attitudes toward science and building foundational problem-solving skills. Research suggests that interdisciplinary learning, especially with technology integration, positively influences students' attitudes toward science (Zhou & Lin, 2024). However, many primary school students struggle with science, often perceiving it as difficult, which has led to a search for more engaging and relevant teaching strategies to improve students' scientific inquiry skills (Ortiz & Aliazas, 2021). Challenges in science education stem from outdated teaching methods (Marcourt et al., 2022), limited exposure to science careers (Kang et al., 2023), and teachers' limited pedagogical content knowledge, all of which hinder students' interest in learning science (Solis-Foronda & Marasigan, 2021).

In Malaysia, data from the Trends in International Mathematics and Science Study (TIMSS) study (1999-2019) reveals a concerning decline in students' science performance, with a drop in the percentage of students meeting advanced benchmarks and average scores falling below the international average (Phang et al., 2021). Despite students' initial interest in science, their confidence wanes over time due to traditional teaching approaches, assessment issues, and limited opportunities to explore science-related careers. To address these challenges, experts recommend curriculum adaptation (DeBarger et al., 2017), pedagogical innovation (Iwuanyanwu, 2019), and the integration of technology into science education, such as the use of games (Yilmaz, 2023).

Game-based learning (GBL) and flipped learning have been identified as promising solutions to these challenges. A meta-analysis by Lei et al. (2022) of 41 studies comparing GBL to traditional methods found that GBL enhances students' understanding of complex scientific concepts, increases engagement, and improves academic outcomes in science learning, particularly in quizzes and exams. Fahdiran et al. (2021) observed improved science learning outcomes in 6th-grade students when the flipped classroom approach incorporated active learning strategies.

Gamification, the integration of game elements into non-gaming contexts, has also gained attention in educational research for its potential to enhance student engagement and motivation (Majdoub & Heilporn, 2024; Rodrigues et al., 2022). Studies suggest that gamification can foster positive learning experiences (Klock et al., 2018) and improve social interaction (Gaonkar et al., 2022). However, gamification is not

without its challenges; for instance, poorly designed gamified activities can lead to frustration, reduced engagement, or a focus on extrinsic rewards rather than intrinsic learning goals (Buckley & Doyle, 2016).

Integrating gamification into flipped learning addresses these challenges by enhancing engagement during out-of-class activities and improving overall flipped learning effectiveness. For instance, Huang et al. (2018) found that gamification improved engagement and performance in a flipped course, while Lo and Hew (2018) reported that gamified flipped learning enhanced academic performance and cognitive engagement compared to traditional approaches. However, challenges such as designing effective gamified components and overcoming technological barriers remain (Gündüz & Akkoyunlu, 2020).

Recent research, including Mulyanti et al. (2023), demonstrates the effectiveness of gamified flipped learning in improving fifth-grade students' science achievement. However, their focus on general science leaves gaps in understanding its impact on specific topics, such as *Machines*, and its applicability in contexts like Malaysia. Additionally, while evidence supports the academic benefits of gamified flipped learning, less is known about its effects on students' learning experiences.

These findings underscore the potential of flipped learning across diverse subjects and settings. However, integrating digital games into this model remains underexplored. This gap highlights the need for further research on how GBL can complement and enhance flipped learning to improve both academic outcomes and student experiences.

2. LITERATURE REVIEW

2.1. Flipped Learning

The concept of flipped learning, introduced by Bergmann and Sams (2012), has transformed traditional teaching methods by reversing classroom activities. Students complete preparatory tasks outside of class and engage in more interactive, teacher-guided activities during in-class sessions. This approach fosters a student-centred learning culture, encouraging deeper engagement with content. In the digital era, flipped learning incorporates synchronous (live, online, real-time) and asynchronous (delayed, not in real-time) learning tasks to maximise flexibility and accessibility.

Recent studies emphasise the adaptability of flipped learning across disciplines and educational levels (see Table 1). For example, Ranoptri et al. (2022) and Erkan and Duran (2023) demonstrated how inquiry-based, and Science, Technology, Engineering, and Mathematics (STEM-integrated) flipped models improved science outcomes and creativity in junior high and primary school students. Similarly, Cheng (2023) found gains in critical thinking, problem-solving, and creativity through STEM-integrated flipped learning. These studies highlight flipped learning's potential to develop higher-order thinking skills. However, the extent to which these benefits are sustained across different contexts requires further exploration.

Flipped learning approaches vary in design and outcomes. Da Silva et al. (2022) and Lee et al. (2021) investigated dialogic and cooperative flipped models, respectively, reporting improvements in science achievement and motivation. While these studies showcase the strengths of flipped learning, they focus

primarily on its structural elements, neglecting the role of motivational strategies that could enhance engagement in pre-class tasks. This gap is particularly concerning given challenges such as the lack of student participation in at-home activities, as noted by Maxwell (2024).

TABLE 1. Critical analysis of research gaps in Flipped Learning

No.	Researchers	Samples	Learning Approaches	Findings
1.	Ranoptri et al. (2022)	7th grade of junior high school students	Inquiry web-based learning multimedia	Improvement of learning outcome in science
2.	Erkan and Duran (2023)	4th grade of primary school students	STEM activities with flipped learning	STEM activities with flipped learning model improved scientific creativity and perceptions
3.	Cheng (2023)	Primary school	STEM activities with flipped learning	Enhancement of critical thinking, problem-solving, and creativity skills
4.	Flores-Gonzales & Flores-Gonzales (2022)	High school students	Virtual learning environment	Enhancement of students' comprehensive learning process in science with increased completion of activities and higher-grade point averages
5.	Da Silva et al. (2022)	High school students	Flipped learning with dialogic strategies	Promotion of student protagonism in science
6.	Lee et al. (2021)	High school students	Cooperative flipped learning (CFL)	CFL showed higher science achievement and motivation than simple flipped learning
7.	Rudoft (2021)	Secondary school students	Flipped learning with interactive video	Significant improvement in understanding mathematical concepts
8.	Fahdiran et al. (2021)	6th grade primary students	Flipped classroom approach with active learning	Improvement of learning outcome in science

2.2. Gamification

Gamification and GBL are increasingly recognised for their potential to enhance student engagement. Unlike GBL, which uses fully immersive games tailored for educational purposes, gamification incorporates game elements such as points, levels, and rewards into traditional educational tasks, fostering motivation without altering the core learning activities (Al-Azawi et al., 2016; Maratou et al., 2023;). This approach has demonstrated success in improving student motivation, engagement, and knowledge retention across various educational contexts (Nipo et al., 2023; Lim et al., 2024). For instance, Nilubol (2023) showed that gamified strategies particularly benefited students with high intrinsic motivation, while Mulyanti et al. (2023) observed enhanced science learning outcomes in fifth-grade students. These findings suggest that gamification could address challenges in flipped learning, such as student disengagement during pre-class tasks, yet successful implementation requires careful alignment of gamified elements with the educational goals (Guerrero-Quinonez et al., 2023).

Yu and Yu's (2023) meta-analysis of gamified flipped classrooms (GFCs) reveals positive effects on academic achievement, motivation, and satisfaction, especially in Asian and African contexts, where

cultural and technological factors amplify the benefits of gamified learning. However, most existing studies have primarily focused on secondary and higher education, leaving a significant gap in understanding the potential of GFCs for primary school students. Integrating gamification into flipped learning combines the principles of flipped classrooms with game design elements to boost student engagement. Gamification addresses the key challenge of engaging students in pre-class activities by providing immediate feedback, rewards, and a sense of progression (Toda et al., 2019). This approach creates a synergy, where flipped learning offers a structured framework for independent and collaborative activities, while gamification enhances engagement and motivation throughout the learning process. Research by Guerrero-Quinonez et al. (2023) and Yu and Yu (2023) highlight the potential of this combined approach, particularly in science education.

Despite its potential, the integration of gamified flipped learning in primary education remains underexplored, as most research treats flipped learning and gamification as separate strategies. This gap is particularly significant in primary education, where gamification has the potential to sustain student motivation and engagement during pre-class activities. The current study seeks to address this gap by examining how gamified flipped learning can enhance primary students' academic achievement and learning experiences, with a specific focus on the science topic of *Machines*. According to Sayeski et al. (2015), the effectiveness of flipped learning is often hindered when students fail to engage with pre-class tasks. By incorporating gamification, this study aims to determine whether such integration can address these challenges and improve the overall flipped learning experience. This investigation is crucial, as existing literature provides limited exploration of gamification as a complementary strategy within the flipped learning model, particularly at the primary school level.

3. THEORETICAL UNDERPINNING OF GAMIFIED FLIPPED LEARNING

The theoretical foundation of gamified flipped learning combines the flipped learning model by Bergmann and Sams (2012) and gamification elements by Toda et al. (2019), offering a framework for integrating gamification into flipped classrooms.

3.1. Foundational pillars of flipped learning

Bergmann and Sams (2012) introduced four foundational pillars of flipped learning: flexible environment, learning culture, intentional content, and professional educator.

- **Flexible environment:** This pillar emphasises both synchronous and asynchronous learning, allowing students to learn at their own pace. In a gamified flipped classroom, this flexibility is enhanced through elements like progress tracking and levels to enable students to navigate their learning journey more dynamically.
- **Learning culture:** The flipped model promotes a student-centred culture, where learners take responsibility for their education. Gamification supports this by encouraging active participation through points, levels, and rewards to reinforce positive behaviours.

- **Intentional content:** Educators select content carefully to suit both independent study and collaborative activities. Gamification helps by breaking content into smaller, engaging segments, often linked to levels or progress indicators to keep students motivated.
- **Professional educator:** Educators play a key role in designing and managing a gamified flipped classroom. They ensure that gamified elements align with learning goals and provide personalised support for students.

By integrating these pillars with gamification, educators can create a more engaging and effective flipped learning environment that addresses the needs of diverse learners.

3.2. Gamification Sub-Elements

Toda et al. (2019) identified several gamification sub-elements that enhance learning engagement and motivation. These elements, when integrated with flipped learning, create a synergistic effect that supports academic achievement and positive learning experiences:

- **Acknowledgement:** Recognising student achievements through badges, trophies, or medals provides extrinsic motivation and encourages active participation.
- **Level:** Offering levels or skill hierarchies motivates students to progress through increasingly challenging tasks, which align with the adaptive nature of flipped learning.
- **Progress:** Progress bars, maps, or steps provide visual feedback on students' advancement, helping them stay on track in both independent and collaborative activities.
- **Points:** Assigning points for completed tasks or correct answers gives students immediate feedback and fosters a sense of accomplishment.
- **Time pressure:** Countdown timers or time-limited challenges encourage focus and efficiency, particularly during in-class activities designed to apply learned concepts.
- **Novelty:** Introducing surprises or unexpected elements within tasks keeps students engaged and prevents monotony, which is an essential aspect of sustaining attention in flipped learning environments.
- **Chance:** Incorporating elements of luck, such as random rewards or outcomes, adds excitement and unpredictability that make the learning experience more engaging.

The integration of Bergmann and Sams' (2012) flipped learning model with the gamification sub-elements proposed by Toda et al. (2019) creates a robust theoretical foundation for gamified flipped learning. This approach not only enhances the structural flexibility and engagement of the flipped classroom but also aligns with modern pedagogical goals, fostering motivation, active learning, and student-centred experiences.

4. RESEARCH CONTEXT, OBJECTIVES, AND QUESTIONS

This study focused on fifth-grade students and investigated the implementation of gamified flipped learning activities to understand their impact on students' academic achievement and learning experiences. The objectives of this study were as follows:

i. To examine the effects of digital gamified flipped learning activities on the academic achievement of Malaysian fifth-grade students.
ii. To explore the learning experiences of fifth-grade students participating in gamified flipped learning activities.

To achieve these objectives, this study addressed the following research questions:

i. What are the effects of digital gamified flipped learning activities on the academic achievement of Malaysian fifth-grade students?
ii. What are the learning experiences of fifth-grade students participating in gamified flipped learning activities?

5. METHOD

5.1. Research Design

This study utilised a mixed-methods approach, employing a quasi-experimental design for the quantitative component and interviews for qualitative data collection. The quasi-experimental design was chosen for its appropriateness in assessing the effectiveness of a digital game-based flipped classroom approach in enhancing student academic achievement on the topic of *Machines*, compared to conventional teaching without flipped learning activities. Specifically, a pre-test and post-test control group design was used. Participants in both the treatment and control groups had similar characteristics but were not randomly assigned due to predetermined group allocations by the school.

To control for potential confounding variables, particularly prior knowledge, a pre-test was administered to both groups before the intervention. The Quade-ANCOVA analysis was then employed to evaluate the effect of the pre-test on post-test scores, using the pre-test scores as a covariate. This approach helped address any pre-existing differences in baseline knowledge between the groups.

This study's data collection process spanned four weeks, comparing a control group and a treatment group in a quasi-experimental design. In Week 1, both groups began with a pre-test to assess their prior knowledge on the science topic of *Machines*. Following the pre-test, the treatment group engaged in five digital games created using Wordwall, which were hosted on a website designed with Google Sites. These games served as the flipped learning component and allowed students to explore the topic asynchronously during Week 2. The control group, meanwhile, did not participate in these pre-class digital activities.

In Week 3, both groups engaged in in-class teaching and learning activities, led by the same teacher to ensure consistency. After completing the in-class sessions, both groups took a post-test in Week 4 to measure knowledge acquisition on the *Machines* topic. In addition, the treatment group participated in interviews to discuss their experiences with the gamified flipped learning activities. Figure 1 illustrates the data collection process involving both the treatment and control groups.

FIGURE 1. Data collection process

Week	Control Group	Treatment Group
	↓	↓
Week 1	Pre-test	Pre-test
	↓	↓
Week 2	↓	**Five Digital Games (Wordwall)**
		↓
Week 3	In-class teaching and learning activities	In-class teaching and learning activities
	↓	↓
Week 4	Post-test	Post-test Interviews

5.2. Samples

For this study, a school in the Kulai district of Malaysia was purposively selected based on the following criteria:

- The school possessed appropriate technological facilities.
- Accessibility for data collection purposes by the researcher.
- The basic background of the involved students could be identified by the researcher.

The study involved a sample of 40 fifth-grade students, selected through cluster random sampling from two classes at a public primary school in the Kulai district. All participants were learning science, focusing on simple machines and their functions. Class 5b was randomly assigned as the treatment group, which experienced the flipped classroom model incorporating digital GBL, while Class 5a, the control group, followed conventional teaching methods. Both groups were taught by the same Science teacher to ensure consistency in instructional quality. To gain deeper insights into the students' experiences with the gamified flipped learning approach, six students from the treatment group were randomly selected for interviews, providing a more thorough understanding of their engagement and learning outcomes.

5.3. Research Instruments

The research instruments included pre- and post-test questions adopted by expert teachers from validated exam questions in the Ministry of Education's Standardised Primary School Evaluation Test, with a focus on the *Machines* topic. These questions assessed students' understanding of the functions of simple machines in various devices. The test consisted of 34 identical subjective questions, each worth one

point, for a total score of 34 points, and was administered during both the pre-test and post-test phases. Based on the Examination Division of the Ministry of Education, the questions were designed to evaluate academic achievement at Bloom's Taxonomy levels of understanding, application, and analysis. To ensure scoring objectivity, all responses were graded by the classroom teacher and verified by an independent teacher using the standardised answer scheme validated by the Ministry, achieving an inter-rater agreement of 82.5%.

While the test items were based on Ministry-approved question banks, which have undergone construct validity checks for the *Machines* topic, additional validation specific to this study's context would have further enhanced their appropriateness for the target age group and topic. A pilot test was conducted with 10 fifth-grade students to check the appropriateness of the test questions, and content validity was established through reviews by two primary science teachers. Additionally, the reliability of the test was confirmed through a test-retest analysis, yielding a high Spearman correlation coefficient of 0.85.

In addition to the test, individual interviews were conducted by the classroom teacher with students in the treatment group to explore their learning experiences during the gamified flipped learning activities. These interviews followed a structured protocol with open-ended questions aimed at capturing students' feelings and perceptions about learning through digital games, their preferences for learning activities, and their views on the effectiveness of the games in supporting their learning. Before data collection, the interview protocol was reviewed and validated by two expert primary school teachers to ensure the questions were appropriate for the fifth-grade level.

5.4. In-class Teaching and Learning Activities

The in-class teaching and learning activities covered the same scope as the five gamified learning activities, including subtopics such as simple machines, complex machines, and the importance of creating tools sustainably. The teaching method used in both groups was a tutorial-based approach within a conventional classroom setting. The same teacher who instructed the treatment group also delivered the content to the control group, ensuring consistency in instructional delivery across both groups. The only difference was that the treatment group engaged with the digital games prior to the in-class activities, providing them with a gamified flipped learning experience before interacting with the conventional instructional content.

5.5. Digital Games

Five digital games were developed for this study using the Analyse, Design, Develop, Implement, and Evaluate (ADDIE) instructional design model (Heinich et al., 2002). The process followed these phases:

1. **Analysis phase:** This phase involved analysing the students' prior knowledge, technological familiarity, and the school's resources, such as computer facilities and internet speed, to ensure the games were compatible.

2. **Design and development phase:** The games were designed using gamification elements like points, badges, and leaderboards (Toda et al., 2019). After development, an Educational Technology expert reviewed the games for alignment with educational goals.
3. **Implementation phase:** The games were pilot-tested with a small group of three fifth-grade students to check their functionality, usability, and engagement. Any technical issues were addressed before the games were used in the actual experiment.
4. **Evaluation phase:** The final phase involved collecting data from students' performance and experiences while using the games in the experiment.

Table 2 presents a detailed mapping of the subtopics, digital game elements, learning activities, and technological tools used in the study.

TABLE 2. Mapping of subtopics, digital game elements, learning activities, and technological tools

Weeks	Subtopics	Gamification Elements (Toda et al., 2019)	Games	Technology Tools
1	Simple Machines 1	• Progress • Point	Digital game 'Speedy Plane'	• Google Sites • Wordwall App
2	Simple Machines 2	• Acknowledgement • Time pressure	Digital game 'Truth Stage'	
3	Complex Machines 3	• Acknowledgement • Novelty	Digital game 'Danger Maze'	
4	Complex Machines 4	• Point • Chance • Novelty	Digital game 'Bomb It'	
5	The Importance of Sustainable Features in Tool Creation	• Point • Time pressure	Digital game 'The Winner'	

Figures 2 and 3 illustrate examples of the Danger Maze and Bomb It games.

FIGURE 2. Danger Maze

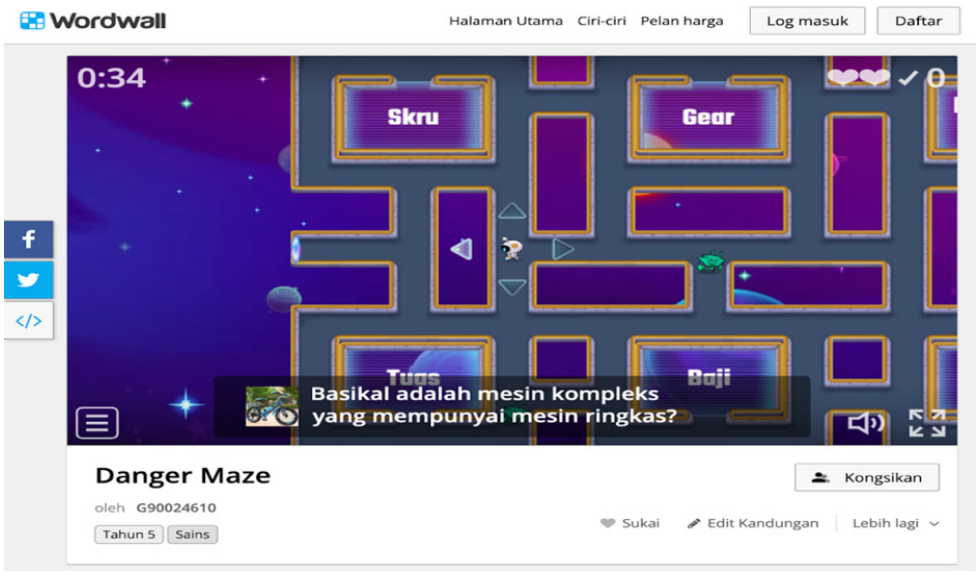
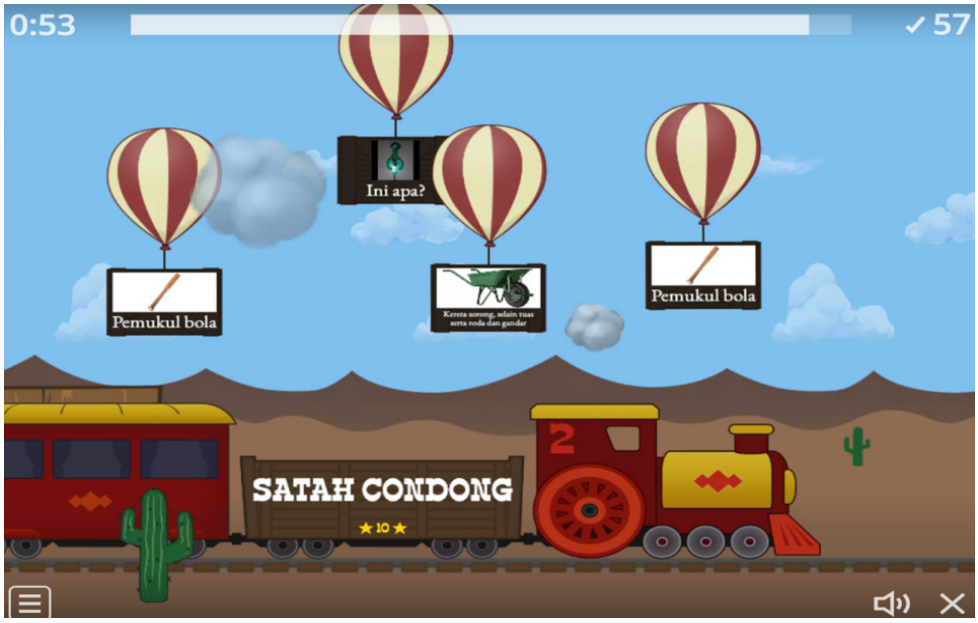


FIGURE 3. Bomb It



5.6. Data Analysis

The data obtained were analysed using descriptive statistical tests, such as the mean and standard deviation, as well as inferential tests, including Quade-ANCOVA. The pre-test was used as a covariate to statistically adjust for initial differences between the treatment and control groups. Given that the data did not follow a normal distribution, Quade-ANCOVA was selected for its robustness against violations of normality assumptions. This method enables comparison of the post-test mean scores between the treatment and control groups while accounting for the effects of the pre-test scores as a covariate.

Additionally, an analysis of student learning experiences in gamified flipped classroom learning was conducted by two expert teachers. This study adapted inductive thematic analysis on the interview data collected, following Braun and Clarke's (2006) six-phase guideline. Since the qualitative data consisted of simple sentences from primary students, themes were identified based on commonalities in the responses provided by the interviewed students. If students expressed similar perspectives regarding the questions asked, those themes were considered stable and significant for highlighting as study findings. The inter-coder reliability between the two teachers was 87.1% agreement.

5.7. Research Ethics

Informed consent was obtained from both the schools and the parents or guardians of the students prior to the study. Participants were assured that their data would remain anonymous, and their privacy would be protected throughout the research process. To ensure that student well-being was not jeopardised, the study was designed with minimal disruption to regular learning activities. Additionally, all

participants were informed that their involvement in the study was voluntary, and they could withdraw at any time without consequence. These measures were taken to uphold ethical standards and ensure the safety and privacy of all students involved.

6. RESULTS

6.1. The Effects of Gamified Flipped Classroom Learning Activities on the Academic Achievement of Malaysian Fifth-Grade Students

Table 3 shows the pre-test and post-test means and standard deviations for both the control and treatment groups. The control group had a pre-test mean of 12.25 with a standard deviation of 2.12, indicating less variability in scores compared to the treatment group, which had a pre-test mean of 9.3 with a standard deviation of 2.57. This suggests that the control group's scores were more consistent before the intervention. After the intervention, both groups showed increased variability in their post-test scores, but the treatment group had a slightly lower standard deviation (3.02) compared to the control group (3.12), indicating more consistency in the treatment group's improvement. The treatment group's post-test mean was 26.9 out of a maximum score of 34, while the control group's post-test mean was 17.

TABLE 3. Descriptive analysis of fifth-grade student's achievement

	Control Group (n = 20)	Treatment Group (n = 20)	Levene's test (p value)	Shapiro-Wilk Test (p value)
Pre-Test Mean	12.25	9.3	.010*	-
Post-Test Mean	17	26.9	-	.041**
Pre-Test Standard deviation	2.12	2.57	-	-
Post-Test Standard deviation	3.12	3.02	-	-

**the homogeneity assumption of the variance did not met*
***post-test data not normally distributed*

The results of Levene's test indicated that the homogeneity of variance assumption was not met for the pre-test scores, with a p-value of 0.010 ($p < 0.05$). This suggests that the variance of the pre-test scores differed significantly between the two groups, thereby violating the assumption of equal variances necessary for standard ANCOVA. As a result, the use of a non-parametric alternative, i.e. Quade-ANCOVA, was considered necessary for this analysis.

Additionally, the Shapiro-Wilk test for normality yielded a p-value of 0.041 for the post-test scores ($p < 0.05$), indicating that the post-test data were not normally distributed. This reinforces the need for a robust statistical method that can handle non-normal data distributions.

Given these violations, Quade-ANCOVA was employed to analyse the significance of the differences in post-test scores between the control and treatment groups. Quade-ANCOVA is a suitable alternative to traditional ANCOVA when the assumption of normality is violated or when variances are unequal between groups, as in this case. Unlike parametric ANCOVA, which assumes normally distributed residuals, Quade-ANCOVA adjusts for covariates while accounting for the non-normality in the data, providing more accurate estimates of treatment effects.

By employing Quade-ANCOVA, the study ensures the validity of its findings despite the violations of homogeneity of variance and normality. This robust approach increases the reliability and validity of the results, making it especially well-suited for educational research, where data often deviate from normality and exhibit heteroscedasticity. Therefore, the use of Quade-ANCOVA enhances the robustness of the study's conclusions while addressing the issues identified through the Levene's and Shapiro-Wilk tests.

Based on the Quade-ANCOVA analysis (see Table 4), a significant difference was observed in the mean post-test scores between the treatment and control groups, with a p-value of less than 0.05. The effect size of the treatment on student achievement was substantial, as indicated by an eta squared value of 0.626.

TABLE 4. The Quade-ANCOVA analysis

Test of Between – Subject Effects						
Dependent Variable: Unstandard Residual						
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Corrected Model	3260.472	1	3260.472	63.737	.000	.626
Intercept	.000	1	.000	.000	1.000	.000
Group	3260.472	1	3260.472	63.737	.000	.626
Error	1943.897	38	51.155			
Total	**5204.368**	**40**				
Corrected Total	**5204.368**	**39**				

*R Squared = .626 (Adjusted R Squared = .617)

6.2. The Learning Experiences of Malaysian Fifth-Grade Students Participating in Gamified Flipped Learning Activities

Based on the thematic analysis of the interview transcripts with students, several learning experiences were identified, as illustrated in Table 5.

TABLE 5. Thematic analysis of interview transcripts

No.	Interview Questions	Example	Theme
1.	How do you feel when you learn through flipped learning with games, what are your feelings?	S1: "I feel so excited when doing quizzes." S2: "Happy, later when answering the exam, that question will be easy because I have studied it." S5: "I feel nervous because games involve competition"	Happy, excited, fun, like it Nervous
2.	Do you like playing digital games on the topic of Machines?	S2: "Like it." S6: "I love it"	Like it
3.	When do you play digital games on the topic of Machines?	S3: "At night." S2: "Whenever I'm free"	Anytime
4.	What is your favourite digital game on the topic of Machines? Why?	S2: "Danger Maze because I always play games like that." S4: "Truth Stage."	Speedy Plane, Danger Maze, Truth Stage
5.	How many times do you play each digital game? If more, why do you play more than once?	S3: "Speedy Plane three times, quiz four times, the rest just once or twice. S5: "More than five times."	Many times (more than five times)
6.	Which option do you prefer, continue learning through digital games or using textbooks as usual in class? Why?	S3: " Digital games made me feel excited." S4: "Digital games because I like it."	Digital games Feel excited
7.	Do you think these digital games help you excel as a student? Why?	S6: "It's easy to learn." S5: "Not sure. I need to play games for other topics"	Easy Unsure

7. DISCUSSION AND CONCLUSIONS

The analysis of between-subject effects reveals a highly significant relationship between participation in gamified flipped learning and conventional learning, with a statistical outcome of $F = 63.737$, $p < .001$ (see Table 4). This significance indicates that the treatment group, engaged in gamified flipped learning activities, experienced notably different learning outcomes compared to the control group. This finding corroborates the studies by Lo and Hew (2018) and Jo et al. (2018), which demonstrated that students in gamified flipped classrooms outperform those in traditional or online independent study settings. Elements of gamification, such as ranking systems and word games, have been shown to increase participation rates, competitive spirit, and interest in online preparation activities (Jo et al., 2018), a trend that is not only evident among higher education students but also among primary school students in this study.

The data analysis further indicates that following the intervention with digital games in flipped learning, students' knowledge and understanding of the topic of *Machines* significantly improved compared to their previous levels. The notable increase in scores among the treatment group relative to the control group underscores the effectiveness of integrating flipped learning with digital games. This finding aligns with Gündüz and Akkoyunlu's (2020) research, which established that gamification in flipped learning environments can enhance achievement.

Additionally, the effect size, measured by partial eta squared (0.626), indicates a large effect (Cohen, 1988), suggesting that 62.6% of the variability in students' learning experiences can be attributed to their participation in gamified flipped learning activities. This substantial effect size underscores the robust impact of these innovative educational approaches on enhancing student engagement, motivation, and comprehension of the subject matter.

The findings also align with the research conducted by Tsay et al. (2018), which reported similar results among second-year business students regarding learning through gamification. This study expands upon Tsay's work by demonstrating that these benefits extend to primary school students, indicating that the effectiveness of the gamification approach is applicable across educational levels. Moreover, learning through digital games can enhance motivation and knowledge, even for students who initially lack experience with such games (Divjak and Tomic, 2011). This study enriches the existing research on gamification by illustrating how the integration of flipped learning enhances its impact. The positive effects observed from the gamified approach can be attributed to the use of flipped learning for the topic of *Machines*.

Students demonstrated remarkable improvement when engaging with gamified flipped learning, as they repeatedly played digital games until achieving their desired scores. They expressed a greater interest in self-paced activities through digital games and consistently allocated time to play games related to the topic of *Machines*. This enthusiasm may stem from the engaging elements of digital games, which rely on factors such as the gaming environment, game design, visual presentation, and mechanical technology (Lacovides et al., 2011). Additionally, students found the provided digital games easy to understand for each concept they learned. Key factors contributing to this effectiveness include perceived ease of use, usefulness, and satisfaction, as identified by Guo et al. (2020). Previous studies have also indicated that well-designed games can be highly playable, enjoyable, and immersive, with most students quickly familiarising themselves with the game elements (Yue & Wan, 2015).

The findings of this study indicate that students' experiences with flipped learning, enhanced by digital games, are overwhelmingly positive (see Table 5). Participants expressed feelings of happiness, excitement, and enjoyment while engaging with this approach. This positivity can be attributed to the student-centred nature of flipped learning, which promotes active participation in the learning process (Siegle, 2013). Furthermore, the integration of digital games as educational tools creates an enjoyable learning environment, challenging students to excel in these interactive activities (Thomas & Mahmud, 2021).

Students interviewed articulated a clear preference for the gamified flipped learning approach over traditional textbook-based learning. They eagerly embraced digital game activities, often engaging repeatedly until they achieved their desired scores. This behaviour reflects their eagerness to grasp additional information and ensure they do not miss key insights during the gamified learning activities. Additionally, students demonstrated high levels of motivation, actively seeking opportunities to participate in digital GBL.

However, one student expressed a neutral response, reporting nervousness due to the competitive nature of the games. For this student, the element of competition, while motivating for many, created a sense of pressure that made the learning experience somewhat stressful. This highlights the importance of designing gamified activities that balance competition with collaboration to ensure all students feel supported in their learning journey.

Using digital games, students effectively recalled relevant content, accurately identifying key elements such as simple machines, wheels and axles, levers, pulleys, screws, inclined planes, wedges, and gears. This mastery was achieved through self-directed learning, as students repeatedly engaged with the digital games until they attained high scores. Beyond the classroom, the flipped learning method not only enhanced students' comprehension of prior knowledge but also reinforced their understanding of new topics through digital games (Du et al., 2014).

In their interviews, students noted that digital games contributed significantly to their academic success and enjoyment of the flipped learning approach. The application of digital GBL methods effectively captured students' attention, promoting increased engagement and boosting their confidence (Tangkui & Tan, 2020). Overall, students' learning experiences with the gamified flipped learning approach were decidedly positive. These findings align with research by Aidoo et al. (2022), which also highlighted students' positive perceptions of the flipped learning method. However, one student expressed uncertainty, stating, "I need to play games for other topics," suggesting a desire for the broader application of digital games to other subjects. This indicates that while the current approach was engaging, there is potential for further expansion to enhance learning across a wider range of topics.

Moreover, this study builds upon previous research, such as the work by Borit and Stangvaltaite-Mouhat (2020), which found that GBL integrated with flipped classrooms resulted in greater enjoyment, engagement, and perceived learning among students in dental education. This study focuses on learning Science at the primary school level, thus expanding the research conducted by Borit and Stangvaltaite-Mouhat (2020).

In conclusion, the integration of GBL and flipped learning presents a highly effective approach to modern education by enhancing student engagement, motivation, and academic performance. Tools such as Wordwall and Google Sites, along with the development of digital games, provided science educators with innovative resources to diversify their teaching methods. The findings revealed significant improvements in fifth-

grade students' academic performance, evidenced by higher post-test scores and positive attitudes toward gamified flipped learning. By combining the interactive and immersive aspects of GBL with the preparatory advantages of flipped learning, students are better equipped with foundational knowledge prior to class and actively engage in deeper learning during in-person sessions. This synergy not only improves comprehension and confidence but also empowers educators with innovative tools to create dynamic and impactful learning experiences, particularly in science education.

7.1. Limitations and future lines of research

Despite the positive findings of this study, several limitations must be acknowledged. First, the research focused specifically on gamified learning related to the topic of *Machines* among fifth-grade students at a single primary school in Malaysia. This narrow scope may limit the generalisability of the results to other contexts or educational levels. Furthermore, the assessment of student improvement in academic achievement was based solely on performance tests, which may not capture the full spectrum of learning outcomes associated with gamified flipped learning, such as critical thinking, creativity, and collaborative skills.

Future studies should consider several recommendations to address these limitations, including conducting research across various educational institutions and regions to assess the effectiveness of gamified flipped learning in different settings and among diverse student populations, which would enhance the generalisability of the findings. Additionally, it would be beneficial to investigate the effectiveness of different gamification elements—competition, reward systems, or specific game mechanics—to identify which elements most effectively enhance student engagement and learning outcomes.

A broader array of assessment methods should be incorporated, including qualitative measures such as observations, student reflections, and peer evaluations alongside quantitative performance tests, to provide a more holistic understanding of the impact of gamified flipped learning on student learning experiences. Furthermore, implementing longitudinal studies could examine the long-term effects of gamified flipped learning on student academic achievement and learning experiences, offering insights into how these methods influence learners over time. Lastly, research should be expanded to include other subjects beyond science, thereby evaluating the applicability and effectiveness of gamified flipped learning across the curriculum.

8. ACKNOWLEDGEMENT

This work was supported/funded by the Ministry of Higher Education (MOHE) Malaysia under Fundamental Research Grant Scheme [FRGS/1/2023/SSI07/UTM/01/2].

9. REFERENCES

Aidoo, B., Tsyawo, J., Quansah, F., & Boateng, S. K. (2022). Students' learning experiences in a flipped classroom: A case study in Ghana. *International Journal of Education and Development using Information and Communication Technology (IJEDICT)*, *18* (1), 67-85.

Al-Azawi, R., Al-Faliti, F., & Al-Blushi, M. (2016). Educational gamification vs. Game based learning: comparative study. *International Journal of Innovation, Management and Technology, 7(4),* 132-136. https://doi.org/10.18178/ijimt.2016.7.4.659

Bergmann, J., & Sams, A. (2012). *Flip your classroom: Reach every student in every class every day.* International Society for Technology in Education.

Borit, M., & Stangvaltaite-Mouhat, L. (2020). *GoDental!* Enhancing flipped classroom experience with game-based learning. *European Journal of Dental Education, 24*(2), 763-772. https://doi.org/10.1111/eje.12566

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Buckley, P., & Doyle, E. (2016). Gamification and student motivation. *Interactive Learning Environments*, *24*, 1162-1175. https://doi.org/10.1080/10494820.2014.964263

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge.

Da Silva, J., Felicio, C. M., & Teodoro, P. V. (2022). Flipped classroom and digital technologies: Didactic possibility for science teaching in an active methodology proposal. *Brazilian Journal of Education*, *Technology, and Society, 17*, 1386-1401. https://doi.org/10.21723/riaee.v17i2.15807

DeBarger, A. H., Penuel, W. R., Moorthy, S., Beauvineau, Y., Kennedy, C. A., & Boscardin, C. K. (2017). Investigating purposeful science curriculum adaptation as a strategy to improve teaching and learning. *Science Teacher Education, 101*(1), 66–98. https://doi.org/10.1002/sce.21249

Divjak, B., & Tomić, D. (2011). The impact of game-based learning on the achievement of learning goals and motivation for learning mathematics-literature review. *Journal of Information and Organizational Sciences*, *35*(1), 15-30. https://jios.foi.hr/index.php/jios/article/view/182

Du, S. C., Fu, Z. T., & Wang, Y. (2014). The flipped classroom—Advantages and challenges. *Proceedings of the 2014 International Conference on Economic Management and Trade Cooperation* (pp. 17–20). Atlantis Press. https://doi.org/10.2991/emtc-14.2014.3

Erkan, H., & Duran, M. (2023). The effects of STEM activities conducted with the flipped learning model on primary school students' scientific creativity, attitudes and perceptions towards STEM. *Science Insights Education Frontiers*, *15*(1), 2175-2225. https://doi.org/10.15354/sief.23.or115

Fahdiran, R., Serevina, V., & Maha Putra, A. V. (2021). The effect of active learning in the flipped classroom learning model on 6th grade science subjects of elementary school. *Journal of Physics: Conference Series*, 1–8. https://doi.org/10.1088/1742-6596/2309/1/012072

Flores-González, E., & Flores-González, N. (2022). The flipped classroom as a tool for learning at High School. *Journal of Critical Pedagogy*, *6*(16), 10–18. https://doi.org/10.35429/JCP.2022.16.6.10.18

Gaonkar, S., Khan, D., & Ashish Singh, M. (2022). Impact of gamification on learning and development. *Journal of Advances in Education and Philosophy*, *6*(2), 63-70. https://doi.org/10.36348/jaep.2022.v06i02.003

Guerrero-Quiñonez, A. J., Quiñónez Guagua, O., & Barrera-Proaño, R. G. (2023). Gamified flipped classroom as a pedagogical strategy in higher education: From a systematic vision. *Ibero-American Journal of Education & Society Research*, *3*(1), 238–243. https://doi.org/10.56183/iberoeds.v3i1.622

Gündüz, A. Y., & Akkoyunlu, B. (2020). Effectiveness of gamification in flipped learning. *SAGE Open*, *10*(4), 1-16. https://doi.org/10.1177/2158244020979837.

Guo, J.-L.; Hsu, H.-P.; Lin, M.-H.; Lin, C.-Y., & Huang, C.-M. (2020). Testing the usability of digital educational games for encouraging smoking cessation. *International Journal Environmental. Research and Public Health*, *17*(8), 2695. https://doi.org/10.3390/ijerph17082695

Heinich, R., Molenda, M., Russell, J. D., & Smaldino, S. (2002). *Instructional media and technologies for learning* (7th ed.). Merrill/Prentice Hall.

Huang, B., Hew, K.F., & Lo, C.K. (2018). Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. *Interactive Learning Environments*, *27*, 1106-1126. https://doi.org/10.1080/10494820.2018.1495653.

Iwuanyanwu, P. N. (2019). What we teach in science, and what learners learn: A gap that needs bridging. *Pedagogical Research, 4*(2), em0032. https://doi.org/10.29333/pr/5780

Jo, J., Jun, H., & Lim H. (2018). A comparative study on gamification of the flipped classroom in engineering education to enhance the effects of learning. *Computer Applications in Engi-*

neering Education, 26(5), 1626-1640. https://doi.org/10.1002/cae.21992

Kang, J., Salonen, A., Tolppanen, S., Scheersoi, A., Hense, J., Rannikmäe, M., Soobard, R., & Keinonen, T. (2023). Effect of embedded careers education in science lessons on students' interest, awareness, and aspirations. *International Journal of Science and Mathematics Education, 21*(1), 211–231. https://doi.org/10.1007/s10763-021-10238-2

Klock, A. C. T., Ogawa, A. N., Gasparini, I., & Pimenta, M. S. (2018). Does gamification matter? A systematic mapping about the evaluation of gamification in educational environments. *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, 2006–2012. https://doi.org/10.1145/3167132.3167347.

Lacovides, I., Aczel, J., Scanlon, E., Taylor, J., & Woods, W. (2011). Motivation, engagement and learning through digital games. *International Journal of Virtual and Personal Learning Environments*, *2*(2), 1–16. https://doi.org/10.4018/jvple.2011040101

Lee, G. G., Jeon, Y. E., & Hong, H. G. (2021). The effects of cooperative flipped learning on science achievement and motivation in high school students. *International Journal of Science Education, 43*(9), 1381–1407. https://doi.org/10.1080/09500693.2021.1917788

Lei, H., Chiu, M. M., Wang, D., Wang, C., & Xie, T. (2022). Effects of Game-Based Learning on Students' Achievement in Science: A Meta-Analysis. *Journal of Educational Computing Research, 60*(6), 1373-1398. https://doi.org/10.1177/07356331211064543

Lim, D., Sanmugam, M., Wan Yahaya, W. A. J., & Al Breiki, M. S. (2024). Redesigning Malaysian university students' player traits from the perspective of game theory: A qualitative study. *International Journal of Information and Education Technology*, *14*(4), 552-558. https://doi.org/10.18178/ijiet.2024.14.4.2076

Lo, C. K., & Hew, K. F. (2018). A comparison of flipped learning with gamification, traditional learning, and online independent study: the effects on students' mathematics achievement and cognitive engagement. *Interactive Learning Environments, 28*(4), 464–481. https://doi.org/10.1080/10494820.2018.1541910

Majdoub, M., & Heilporn, G. (2024). How does gamified L2 learning enhance motivation and engagement: A literature review. In Z. Çetin Köroğlu & A. Çakır (Eds.), Fostering Foreign Language Teaching and Learning Environments With Contemporary Technologies (pp. 134-173). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3693-0353-5.ch007

Maratou, V., Ennami, F., Luz, F., Abdullahi, Y., Medeisiene, R. A., Sciukauske, I., Chaliampalias, R., Kameas, A. D., Sousa, C., & Rye, S. (2023). Game-based learning in higher education using analogue games. *International Journal of Film and Media Arts, 8*(1), 68–84. https://doi.org/10.24140/ijfma.v8.n1.04

Marcourt, S. R., Aboagye, E., Armoh, E. K., Dougblor, V. V., & Ossei-Anto, T. A. (2022). Teaching method as a critical issue in science education in Ghana. *Social Education Research, 4*(1), 82–90. https://doi.org/10.37256/ser.4120232058

Maxwell, M. (2024). *The effects of a gamified flipped classroom on first-generation low income student motivation and achievement in a Georgia High School Mathematics Class*. [Tesis de doctorado, Georgia State University]. https://doi.org/10.57709/36975836

Mulyanti, F., Abidin, Y. D., & Suharto, N. (2023). Pengaruh flipped classroom berbasis gamifikasi dan motivasi belajar terhadap hasil belajar IPA. *Pendas: Jurnal Ilmiah Pendidikan Dasar*, *8*(1), 4207-4219. https://doi.org/10.23969/jp.v8i1.7594

Nilubol, K. (2023). The feasibility of an innovative gamified flipped classroom application for university students in EFL context: An account of autonomous learning. *English Language Teaching, 16*(8), 24-38. https://doi.org/10.5539/elt.v16n8p24

Nipo, D., Gadelha, D., Da Silva, M., & Lopes, A. (2023). Game-based learning: possibilities of an instrumental approach to the FEZ game for the teaching of the Orthographic Drawings System Concepts. *Journal on Interactive Systems, 14*(1), 1–16. https://doi.org/10.5753/jis.2023.3190

Ortiz, A. M. L., & Aliazas, J. V. C. (2021). Multimodal representations strategies in teaching science towards enhancing scientific inquiry skills among grade 4. *International Multidisciplinary Research Journal, 3*(3), 107–118. https://doi.org/10.54476/iimrj241

Phang F. A., Khamis, N., Nawi, N. D., & Pusppanathan, J. (2021). TIMSS 2019 Science Grade 8: Where is Malaysia standing? *Asean Journal of Engineering Education, 4*(2), 37-43. https://doi.org/10.11113/ajee2020.4n2.10

Ranoptri, D., Mustaji, M. , & Syaiful Bachri, B. (2022) Development of web bases inquiry learning with the flipped classroom model in science learning for 7th grade of junior high school. *Prisma Sains: Journal Pengkajian Ilmu dan Pembelajaran MIPA IKIP Mataram*, *10*(2), 316-326. https://doi.org/10.33394/j-ps.v10i2.4942

Rodrigues, L., Arndt, D., Palomino, P., Toda, A., Klock, A., Avila-Santos, A., & Isotani, S. (2022). Affective memory in gamified learning: A usability study. *Anais do XXXIII Simpósio Brasileiro de Informática na Educação* (pp. 585-596). https://doi.org/10.5753/sbie.2022.225748

Rudolf, E. (2021). Student as a performer of flipped learning in mathematics upgraded with interactive video. *Natural Science Education*, *18*(1), 22-28. https://doi.org/10.48127/gunse/21.18.22

Sayeski, K. L., Hamilton-Jones, B., & Oh, S. (2015). The Efficacy of IRIS STAR Legacy Modules under different instructional conditions. *Teacher Education and Special Education, 38*(4), 291–305. https://doi.org/10.1177/0888406415600770

Siegle, D. (2013). Technology: Differentiating instruction by flipping the classroom. *Gifted Child Today, 37*(1), 51–55. https://doi.org/10.1177/1076217513497579

Solis-Foronda, M., & Marasigan, A. C. (2021). Understanding the students' adversities in the science classroom. *Journal of Education and E-Learning Research, 8*(1), 52–58. https://doi.org/10.20448/journal.509.2021.81.52.58

Tangkui, R., & Tan, C. K. (2020). Digital game-based learning using minecraft: enhacement of pupils' achievement in fractions? *E-Jurnal Penyelidikan dan Inovasi*, *7*, 75–90.

Thomas, D. S., & Mahmud, M. S. (2021). Digital game-based learning in Mathematics teaching: a literature review. *Jurnal Dunia Pendidikan, 3*(3), 158-168.

Toda, A.M., Klock, A.C.T., Oliveira, W., Palomino, P. T., Rodrigues, L., Shi, L., Bittencourt, I., Gasparini, I., Isotani, S., & Cristea, A. I. (2019). Analysing gamification elements in educational environments using an existing Gamification taxonomy. *Smart Learning Environments*, *6*(16), 1-14. https://doi.org/10.1186/s40561-019-0106-1

Tsay, C. H. H., Kofinas, A., & Luo, J. (2018). Enhancing student learning experience with technology-mediated gamification: An empirical study. *Computers & Education*, *121*, 1-17. https://doi.org/10.1016/j.compedu.2018.01.009

Cheng, X. (2023). Flipped learning model: An effective approach to primary school STEM education. *Science insights education frontiers*, *15*(1), 2145-2146. https://doi.org/10.15354/sief.23.co044

Yılmaz, Ö. (2023). The role of technology in modern science education. In Ö. Baltacı (Ed.), *Current research in education - VI* (pp. 35-60). Özgür Publications. https://doi.org/10.58830/ozgur.pub383.c1704

Yu, Q., & Yu, K. (2023). The effects of gamified flipped classroom on student learning: evidence from a meta-analysis. *Interactive Learning Environments*, *32*(9), 1–16. https://doi.org/10.1080/10494820.2023.2209791

Yue, W. S., & Wan, W. L. (2015). The effectiveness of digital game for introductory programming concepts. *Proceedings of the 10th International Conference for Internet Technology and Secured Transactions (ICITST)* (pp. 421–425). https://doi.org/10.1109/ICITST.2015.7412134

Zhou, X., & Lin, X. (2024). A Survey and study on interdisciplinary learning attitudes of primary school students under the STEM education concept. *Journal of Higher Vocational Education, 1*(3), 149–155. https://doi.org/10.62517/jhve.202416325

Innoeduca. International Journal of Technology and Educational Innovation
Vol. 11. No. 1. June 2025 - pp. 169-183 - ISSN: 2444-2925 DOI: https://doi.org/10.24310/ijtei.111.2025.20318

Calidad e innovación del profesorado en educación superior. Validación del cuestionario de opinión mediante ecuaciones estructurales

Quality and innovation of teaching staff in higher education: Validation of the opinion questionnaire using structural equation modeling

RECIBIDO 22/07/2024 ACEPTADO 16/12/2024 PUBLICADO 01/06/2025

Rafael Barea Colorado
ORCID 0009-0003-9910-2495
Centro de Estudios Universitarios Cardenal Spínola CEU (Universidad de Sevilla), España
rbarea@ceu.es

Carmen Llorente Cejudo
ORCID 0000-0002-4281-928X
Universidad de Sevilla, España
karen@us.es

Beatriz Hoster Cabo
ORCID 0000-0002-7792-7870
Centro de Estudios Universitarios Cardenal Spínola CEU (Universidad de Sevilla), España
bhoster@ceu.es

RESUMEN

La calidad y la innovación en las instituciones educativas, así como la calidad docente del profesorado, se ha convertido en un elemento esencial y de análisis en la sociedad del siglo XXI por el impacto que tiene el factor docente en los resultados derivados del proceso de enseñanza-aprendizaje. El objetivo de este trabajo se centra en analizar el grado de fiabilidad y validez del Cuestionario de Opinión del Alumnado sobre la Actuación Docente del Profesorado. Este cuestionario ha sido diseñado, de forma estandarizada, por la Oficina General de Calidad de la Universidad de Sevilla, y es utilizado para medir aspectos que tienen que ver con la calidad docente del profesorado en el aula, en todas las titulaciones universitarias de dicha institución. El cuestionario fue administrado a una muestra de 6.692 estudiantes pertenecientes a cinco titulaciones universitarias diferentes. Se emplearon técnicas de análisis factorial exploratorio (AFE) y análisis factorial confirmatorio (AFC), mediante la utilización de ecuaciones estructurales. Los resultados del estudio demuestran, por un lado, la unidimensionalidad del instrumento, así como el alto grado de fiabilidad y validez de este. Esto nos lleva a poder ratificar que el cuestionario empleado es un instrumento válido y fiable para medir la actuación docente del profesorado, teniendo en cuenta la opinión del alumnado. Los resultados obtenidos permiten corroborar el objetivo del estudio de forma que la herramienta empleada para valorar el grado de calidad de la actuación docente del profesorado es adecuada para poder medir la satisfacción docente del profesorado en base a la opinión que los alumnados tienen de la docencia impartida en el aula.

PALABRAS CLAVE Evaluación docente; Validación; Ecuaciones estructurales; Cuestionarios; Calidad docente universitaria.

ABSTRACT

Quality and innovation in educational institutions, as well as the teaching quality of faculty members, have become essential elements of analysis in 21st-century society due to the impact of the teaching factor on the outcomes of the teaching-learning process. The objective of this study is to analyze the reliability and validity of the Student Opinion Questionnaire on Faculty Teaching Performance. This questionnaire was standardized and designed by the General Quality Office of the University of Seville and is used to assess aspects related to the teaching quality of faculty members in the classroom across all university programs at the institution. The questionnaire was administered to a sample of 6,692 students from five different academic programs. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) techniques were employed using structural equation modeling. The results of the study demonstrate, on the one hand, the one-dimensionality of the instrument, as well as its high degree of reliability and validity. This allows us to confirm that the questionnaire used is a valid and reliable tool for measuring faculty teaching performance, based on students' opinions. The results obtained corroborate the study's objective, confirming that the tool used to evaluate the quality of faculty teaching performance is appropriate for assessing teaching satisfaction based on students' perceptions of classroom teaching.

KEYWORDS Teaching evaluation; Validation; Structural equations; Questionnaires; University teaching quality.

1. INTRODUCCIÓN

La calidad y la innovación en las instituciones educativas, en líneas generales y, en particular, la calidad docente del profesorado se ha convertido en un elemento esencial y de análisis en la sociedad del siglo XXI. En este sentido, existen ya diferentes estudios (Ajjawi y Boud, 2017; Aláez et al., 2018; Delgado-López-Cózar, 2022; Elitok y Cavus, 2019; Navarro y Ramírez, 2018) que abordan y respaldan el tema de la importancia de la evaluación institucional de la calidad docente del profesorado universitario como factor fundamental del proceso de aprendizaje del estudiante. A la hora de analizar la calidad del profesorado, hay dos cuestiones fundamentales: el acceso a la profesión docente (ingreso a la formación inicial y su currículo, e incorporación al sistema educativo) y las políticas de formación permanente (Campos Soto et al., 2020; Urkidi et al, 2020; Walton-Fisette y Sutherland, 2020).

En los últimos años, el proceso de habilitación y evaluación del profesorado universitario en España ha evolucionado significativamente, especialmente tras la aprobación del Real Decreto 678/2023, en el que se unifica y simplifica el sistema de acreditación estatal y los concursos de acceso a plazas docentes. Este nuevo marco normativo, alineado con la Ley Orgánica del Sistema Universitario (LOSU), prioriza la evaluación integral de la actividad académica mediante métodos cualitativos y un uso responsable de indicadores cuantitativos. Además, fomenta la inclusión de aspectos como la ciencia abierta y la innovación docente, adaptándose a estándares internacionales para mejorar la calidad del sistema universitario (García-Herrero et al., 2024; Ministerio de Universidades, 2023).

La calidad de la docencia impartida ha adquirido un papel central en estos procesos. Por ejemplo, las universidades integran las encuestas de satisfacción estudiantil como herramienta principal para evaluar aspectos como la planificación, desarrollo y resultados de la actividad docente. Estas encuestas, digitalizadas desde 2018 en algunas instituciones, proporcionan un análisis detallado que influye directamente en la evaluación y mejora continua del profesorado (Universidad de La Rioja, 2024). Paralelamente, el programa ACADEMIA de ANECA implementa una nueva metodología para valorar tanto la trayectoria

investigadora como las contribuciones docentes, exigiendo criterios específicos de calidad y relevancia académica (ANECA, 2024).

Estas reformas buscan no solo una mayor transparencia y eficiencia en los procesos de selección y promoción, sino también alinear la producción académica con las necesidades sociales y el desarrollo sostenible, respondiendo a las demandas de un entorno académico cada vez más competitivo y diversificado. Esto refleja un compromiso por integrar la excelencia docente y la innovación en el núcleo de la evaluación del profesorado universitario.

Dentro de los estándares de calidad, uno de especial importancia es el referente a la función docente del profesorado en todo lo que respecta al desempeño de su labor de enseñanza, debiendo siempre tener muy en cuenta que quien ejerce la docencia universitaria es un profesional, pero no necesariamente un profesional de la pedagogía y la didáctica. No es una persona que ha estudiado la carrera de docente. No obstante, cuando este es evaluado, se tienen en cuenta estas dos grandes dimensiones: dominio de los conocimientos que imparte, que tiene que ver con su formación profesional; y dominio de la enseñanza, entendida esta como su aplicación metodológica. Este debe ser un especialista en pedagogía y didáctica, a fin de poder lograr en sus estudiantes nuevos aprendizajes, o identificar las necesidades del alumnado, dando respuesta a las demandas formativas y competenciales que la sociedad actual exige (Montero et al., 2017; Rodríguez et al., 2020). Para lograr enseñanzas de calidad, una de las piezas clave es la profesionalidad y experticia del profesorado universitario que juega un importante papel en el proceso educativo. Sin embargo, en la denominada era digital, el rol del profesor universitario se vuelve complejo, ya que implica la importante tarea de crear las condiciones adecuadas dentro del proceso de enseñanza-aprendizaje, así como de poseer las competencias y habilidades necesarias para ejercer su profesión con eficacia. Esto solo puede lograrse a través de la formación continua que reciban los profesores, la cual les permitirá adquirir y aplicar un conjunto de conocimientos disciplinares y pedagógicos, así como habilidades tecnológicas y didácticas para llevar a cabo su labor docente con excelencia y fortalecer su labor académica (Cervantes Mata, 2020; Rivero Castro y Medeiros Urzua, 2023). Por tanto, es un motivo personal para que desde los centros educativos puedan adecuar la actuación docente del profesorado y por consiguiente su satisfacción en el alumnado en función del grado de uso didáctico y seguro que damos este tipo de dispositivos. Para conseguirlo, se requiere de un profesorado cualificado para abordar el proceso de enseñanza-aprendizaje de los discentes con garantías y, es aquí, donde radica la importancia de conocer más a fondo qué es la Competencia Digital Docente (CDD), y algunos de los diferentes modelos que se han desarrollado para su formación.

En este sentido, y como consecuencia de las adaptaciones realizadas en los planes de estudio al Espacio Europeo de Educación Superior (EEES), se ve necesario llevar a cabo una serie de modificaciones en ellos que requieren de una mejora de la calidad en cuanto a lo que al desarrollo de sus actividades se refiere. Para ello, se ha de utilizar una serie de mecanismos de carácter sistemático que permitan apoyar los procesos de innovación y desarrollo científico-técnico. Estos suponen una serie de cambios a nivel organizativo y pedagógico, pudiendo así favorecer el interés hacia determinados métodos pedagógicos innovadores (Marqués, 2008). En todo este proceso de cambio, se piensa que la valoración que los estudiantes puedan realizar sobre el grado de satisfacción con la calidad docente de la enseñanza universitaria puede resultar un factor esencial que permitiría emprender acciones de mejora encaminadas a efectuar una inversión en reforzar los procesos de calidad docente a través de una serie de acciones formativas que la completen. Para

ello, es necesario contar con instrumentos capaces de medir una serie de criterios que vendrán a justificar el análisis de la calidad docente, a través de indicadores tales como: el buen trato en la relación docente-estudiantes, el dominio temático de la materia que se imparte, el método de enseñanza, la puntualidad y la asistencia, y la manera de evaluar los aprendizajes (García Garduño y Medécigo Shej, 2014).

Con ese mismo propósito, Muñoz et al. (2002) realizaron una revisión sobre los distintos instrumentos y dimensiones que se estaban utilizando en universidades con el objetivo de poder diseñar una herramienta de 40 indicadores, distribuidos en 10 dimensiones (cumplimiento con las obligaciones, infraestructuras, programa, conocimiento, interrelación con la materia, metodología, materiales, actitud del docente, evaluación, prácticas y satisfacción de los colectivos implicados). En la misma línea, Molero López-Barajas y Ruiz Carrascosa (2005) diseñan un cuestionario de evaluación compuesto por 25 ítems agrupados en tan solo cuatro dimensiones (Interacción con el alumnado, Metodología, Obligaciones docentes y evaluación, Medios y recursos), para evaluar la docencia universitaria en la Universidad de Jaén. Posteriormente, González López y López Cobo (2010) a través de su estudio, tratan de establecer la tipología de competencias docentes que las universidades españolas consideran vitales a la hora de definir el rol de un eficaz docente universitario, y establecen así un sistema de 40 indicadores destinados a evaluar el grado de calidad docente universitaria.

Posteriormente, autores como Pozo Muñoz et al. (2011) realizaron un estudio consistente en la realización de una revisión por comunidades autónomas en las que se llevaba a cabo la evaluación de la actividad docente del profesorado a través del modelo DOCENTIA, donde explicaron la adaptación realizada por Pozo Muñoz et al. (2009) en la universidad de Almería con la única finalidad de crear un marco común de evaluación del profesorado que pudiera servir de referencia para ser utilizado en el resto de universidades andaluzas, introduciendo en este modelo aspectos de innovación y mejora que pudieran contribuir a dotar de mayor calidad la enseñanza universitaria.

Por ello, y a pesar de las distintas evaluaciones que a lo largo de los tiempos se han venido utilizando, y cuyas finalidades han sido diferentes y variadas teniendo como origen instituciones diversas, señalar que todas han tenido como marco central al estudiante, tomándolo como referencia y punto de partida en todo lo concerniente a la evaluación del profesorado (Clayson, 2009; García Berro et al., 2011; González López y López Cobo, 2010; Harvey, 2003; Jiménez et al., 1999; Mat Dau y Abu Kassim, 2011).

Llevar a cabo una evaluación del profesorado y, más concretamente, la evaluación de la calidad docente de la enseñanza implica adoptar una técnica para su análisis. En este sentido, son diferentes los estudios que avalan que la técnica más empleada para realizar esta práctica es el cuestionario (López Cámara et al., 2015). Los defensores de esta propuesta vienen a argumentar que el alumnado es el más fiel observador de la enseñanza (Tejedor y García-Valcárcel, 1996), son los que reciben la enseñanza y, por tanto, los que van a conocer de primera mano qué es lo que ocurre dentro del aula (Álvarez et al., 1999; Mateo, 2000), así como que la calidad de la docencia impartida en el aula está íntimamente relacionada con la evaluación que los estudiantes realizan de esta (De-Juanas Oliva y Beltrán Llera, 2013).

Aunque existen universidades que se apoyan en un mismo modelo de evaluación de la docencia, no debemos olvidar que la evaluación de esta debe ser concebida como un proceso abierto, holístico, en continuo cambio y en perfeccionamiento; es decir, un proceso formativo adaptado al contexto en el que se va a desarrollar. A pesar de que, tras muchas revisiones no existe consenso sobre cuáles son los mejores indicadores para medir el impacto de la calidad docente del profesorado (Dorta González y Dorta González, 2012), lo que

sí parece que está claro es que las encuestas utilizadas para medir la opinión del alumnado, en cuanto al grado de calidad docente universitaria, pueden tener un efecto positivo y no coercitivo, siendo beneficiosas en la actividad docente del profesorado (García Berro et al., 2011).

Teniendo en cuenta todo lo anteriormente reseñado, y sabiendo que el cuestionario es considerado como un instrumento que pretende cuantificar características en las que no hay respuestas correctas o incorrectas, como ocurre en la medición de atributos relacionados con actitudes, opiniones, intereses o disposiciones (Levy y Varela, 2003), este estudio se ha centrado en medir la fiabilidad y la validez de un instrumento destinado a establecer el grado de satisfacción del alumnado con la actuación docente del profesorado universitario, o lo que es lo mismo, el nivel de calidad de la enseñanza impartida por el docente en el aula universitaria.

2. MATERIAL Y MÉTODO

Se plantea el objetivo principal del estudio en base a: Validar y analizar la fiabilidad del instrumento de evaluación "Cuestionario de Opinión del Alumnado sobre la Actuación Docente del Profesorado" (ver Anexo I).

2.1. Población y Muestra

Para la selección de la muestra se optó por un muestreo no probabilístico o dirigido, en donde la elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características de la investigación o los propósitos del investigador (Hernández-Sampieri et al., 2014).

Con 6.679 participantes en una población de 6.692 alumnos y alumnas, se asegura el 99,8% de dicha población. Conjuntamente, se obtiene un margen de error del <1%, con un nivel de confianza del 95%.

El estudio toma como población objeto de análisis los 6.692 estudiantes a los que se tuvo acceso, de los cuales, solo 6.679 respondieron íntegramente el cuestionario del estudio, con la siguiente distribución: 2.371 (35,5%) pertenecían a los estudios de Grado en Educación Primaria, 2.157 (32,3%) a los de Infantil, 886 (13,3%) a los de Ciencias de la Actividad Física y del Deporte, 456 (6,8%) a Derecho, y 806 (12,1%) al Máster Universitario en Profesorado de Enseñanza Secundaria Obligatoria y Bachillerato, Formación Profesional y Enseñanza de Idiomas (MAES). Hay que apuntar que, de los 6.679 que respondieron el cuestionario, 1.193 (20,3%) eran hombres y 4.686 (79,7%) mujeres, con una media de edad de entre 18 y 23 años.

En cuanto al perfil de los alumnos seleccionados nos centraremos para definirlo en dos criterios significativos. Por un lado, la asistencia a clase, donde la gran mayoría del alumnado afirma asistir a clase más del 75% de las horas lectivas (89%), seguido de los que acuden a las aulas entre un 51-75% (7,1%) y, por último, los que van entre un 26-50% (2,1%) y menos del 25% (1,8%). Como segundo criterio, se optó por el interés que los participantes manifiestan hacia las asignaturas, resultando bastante (40,3%), seguido de los que les interesan mucho (33,4%) y, por último, aquellos cuyo interés por estas es algo (22,6%) o nada (3,7%).

2.2. Instrumento de recogida de datos

Para la obtención de los datos se ha utilizado el instrumento denominado Cuestionario de Opinión del Alumnado sobre la Actuación Docente del Profesorado (Anexo I). Cada ítem se mide en una escala Likert de 5

intervalos, donde 1 representa el valor mínimo de acuerdo y 5 el máximo. Además, existe la opción de No Sabe "NS" destinada a cuando un sujeto no tiene suficiente información sobre el enunciado o considera que no procede responder. Dicha herramienta viene definida en el manual de calidad de la Universidad, y es de obligada aplicación para todas las titulaciones que se imparten en la misma, quedando exenta de modificación y adaptación por parte de los centros o titulaciones.

El cuestionario consta de dos partes, una primera parte destinada a recoger datos tales como profesor, título y asignatura a evaluar y aspectos sociodemográficos de los participantes: género, edad, curso más alto y bajo en el que se encuentra matriculado, interés por la asignatura, grado de dificultad y porcentaje de asistencia a clase y tutorías del docente objeto de evaluación. Dichos datos pueden utilizarse para una descripción de la muestra que permita estudiar su posible influencia sobre otras variables del cuestionario. Esto se complementa con la segunda parte del cuestionario, en la que se presentan los 18 ítems para evaluar la actuación docente del profesorado, en base a la opinión del alumnado (tabla 1), junto a una pequeña introducción en la que se describen las premisas esenciales y las indicaciones de cumplimentación del cuestionario.

Los ítems se encuentran agrupados bajo una única dimensión, en la que de forma descriptiva y secuencial se encuentran enunciados, sin seguir ningún tipo de orden, todos los aspectos que hacen referencia a la óptima actuación docente del profesorado en el aula.

TABLA 1. Descripción de los ítems del cuestionario

Dimensión	Ítems	Descriptor	M (μ)	DT (σ)
Opinión del alumnado sobre la actuación docente del profesorado	1	Me ha dado orientaciones para conocer el proyecto docente de la asignatura	4.06	.768
	2	Su docencia se ajusta a la planificación prevista en el proyecto docente	4.45	.675
	3	Me atiende adecuadamente en tutorías	4.31	.702
	4	Su horario de tutorías es adecuado	4.49	.670
	5	La bibliografía y demás material docente recomendado me están resultando útiles para el seguimiento de la asignatura	4.16	.794
	6	Su docencia está bien organizada	4.16	.731
	7	Los medios que utiliza para impartir su docencia son adecuados	4.32	.675
	8	La bibliografía y demás material docente recomendado están a disposición de los estudiantes	4.44	.997
	9	Explica con claridad	3.94	.974
	10	Se interesa por el grado de comprensión de sus explicaciones	3.61	.997
	11	Expone ejemplos para poner en práctica los contenidos de la asignatura	3.67	.968
	12	Resuelven las dudas que se plantean	4.05	.763
	13	Fomenta un clima de trabajo y participación	4.16	.794
	14	Motiva a los/las estudiantes para que se interesen por la asignatura	4.44	.731
	15	Trata con respeto a los/las estudiantes	4.48	.620
	16	Su docencia me está ayudando a alcanzar los objetivos de la asignatura	3.92	.976
	17	Los criterios y sistemas de evaluación me parecen adecuados	4.48	.615
	18	En general, estoy satisfecho/a con la actuación docente desarrollada por este/a profesor/a	4.43	.701

2.3. Procedimiento de recogida y análisis de datos

El cuestionario se ha administrado en formato digital, a través de la plataforma de calidad de la que dispone el centro universitario en el que se ha llevado a cabo el estudio. El acceso a la misma por parte del alumnado se lleva a cabo a través de un enlace web o de código Qr que se les ha proporcionado por el evaluador. La aplicación del cuestionario se lleva a cabo en dos momentos a lo largo del curso, coincidiendo con la finalización de la docencia en cada uno de los cuatrimestres, y anterior a la realización de las pruebas de evaluación correspondiente a la convocatoria final, de forma que los resultados no se vean influidos por las calificaciones obtenidas en cada una de las materias/profesores objeto de evaluación. En cada una de las administraciones, por parte del evaluador, se le explicó al alumnado la finalidad de esta y se le pidió su colaboración de forma veraz. En todo momento se ha asegurado el anonimato de todos los participantes y la adecuada custodia de los resultados.

La fiabilidad, validez discriminante y validez convergente del cuestionario se han calculado aplicando los coeficientes: alfa de Cronbach, omega de McDonald, fiabilidad compuesta (CR), varianza media extractada (AVE) y varianza máxima compartida (MSV). Complementariamente, y con la finalidad de comparar estos resultados, se ha utilizado el método de análisis inferencial entre ítems y dimensiones. Para ello, se ha empleado la técnica de análisis correlacional bivariado por medio del coeficiente de correlación p de Spearman. A su vez, la validez del constructo de la prueba se ha obtenido mediante un análisis factorial exploratorio (AFE), utilizando el método de componentes principales. A su vez, debemos tener en cuenta que se ha utilizado un análisis factorial exploratorio (AFE) con idea de determinar la validez del constructo de la prueba. El método utilizado para la selección de los factores es el método de componentes principales. Los factores obtenidos son rotados utilizando el método de Varimax con normalización Kaiser. Una vez extraído el número de factores, utilizaremos un análisis factorial confirmatorio (AFC), con idea de "confirmar" mediante el análisis de la muestra las relaciones propuestas a partir de la teoría explicativa que se haya decidido utilizar como referencia y a través del cual comprobaremos si las medidas teóricas del modelo son consistentes a través del modelado de diagramas y uso de ecuaciones estructurales (Ruiz Díaz et al., 2010), o lo que es lo mismo, vendremos a verificar si los datos obtenidos se ajustan al modelo de medición hipotético planteado a través del análisis factorial exploratorio. Teniendo en cuenta lo anterior y el concepto de "ajuste", procedemos a utilizar el modelo de mínimos cuadrados ponderados (WLS), con idea de proporcionar las estimaciones consistentes en muestras que no se ajustan a los criterios de normalidad (Ruiz Díaz et al., 2010). Para ello hemos optado por utilizar el programa informático Amos, capaz de establecer relaciones complejas hipotéticas entre variables, usando un modelo de ecuaciones estructurales (SEM). Simultáneamente, se ha procedido a comprobar que los datos no se distribuyen normalmente a través de un estudio descriptivo en el que se ha tenido en cuenta la asimetría y la curtosis.

3. RESULTADOS

La fiabilidad del Cuestionario de Opinión del Alumnado sobre la Actuación Docente del Profesorado se ha calculado a través del coeficiente de alfa de Cronbach de forma global, para su dimensión. Los resultados muestran un índice de alfa de Cronbach de .9 para el cuestionario.

Tanto para la prueba de alfa de Cronbach como para la omega de McDonald se establece un índice de fiabilidad bastante alto (>.988), lo que hace indicar que el cuestionario posee un alto índice de fiabilidad (O´Dwyer y Bernauer, 2014) ya que sus valores superan .7 (Lévy Mangin et al., 2006). En la tabla 2 se presenta el índice de fiabilidad del cuestionario para la dimensión establecida. Conjuntamente a todo esto, se calculan los coeficientes de fiabilidad compuesta (CR), varianza media extractada (AVE) y varianza máxima compartida (MSV), aunque no lo utilizaremos de mucho de nuestro estudio, ya que nuestro cuestionario es unidimensional. Tanto los resultados como los valores de referencia tomados para el ajuste del modelo se muestran en la tabla 2 (Hair et al., 2010).

TABLA 2. Validez convergente y discriminante del modelo

Dimensión	CR	Ajuste	AVE	Ajuste	MSV	Ajuste
A	.988	CR>.7	.829	AVE>.5	.658	MSV<AVE

Fuente: elaboración propia

Todas las cifras obtenidas son los valores de referencia. Por tanto, y aunque el cuestionario sea unidimensional, estos datos nos hacen ver la fiabilidad el modelo (CR), así como su validez convergente (AVE) y discriminante (MSV). A continuación, se procede a realizar el análisis correlacional simple de cada ítem, en la única dimensión del cuestionario. Los resultados quedan expuestos en la tabla 3.

TABLA 3. Correlación de los ítems en la dimensión asociada del cuestionario

Ítems	Valor Dimensión
1	.81
2	.83
3	.83
4	.81
5	.84
6	.86
7	.88
8	.81
9	.88
10	.89
11	.87
12	.87
13	.87
14	.88
15	.69
16	.91
17	.85
18	.92

Fuente: elaboración propia

Todos los valores extraídos son superiores a .81, excepto el correspondiente al ítem 15, que expresa un valor de .69. Aun así, todos los ítems se encuentran integrados en su única dimensión (Carmines y Zeller, 1979).

Estos resultados posteriormente fueron contrastados a partir del coeficiente omega de McDonald, calculados a partir de los pesos factoriales de la matriz de componente rotado de forma global. El resultado obtenido en dicha prueba viene a corroborar la fiabilidad obtenida con anterioridad, dando un resultado de .988 en el coeficiente omega de McDonald.

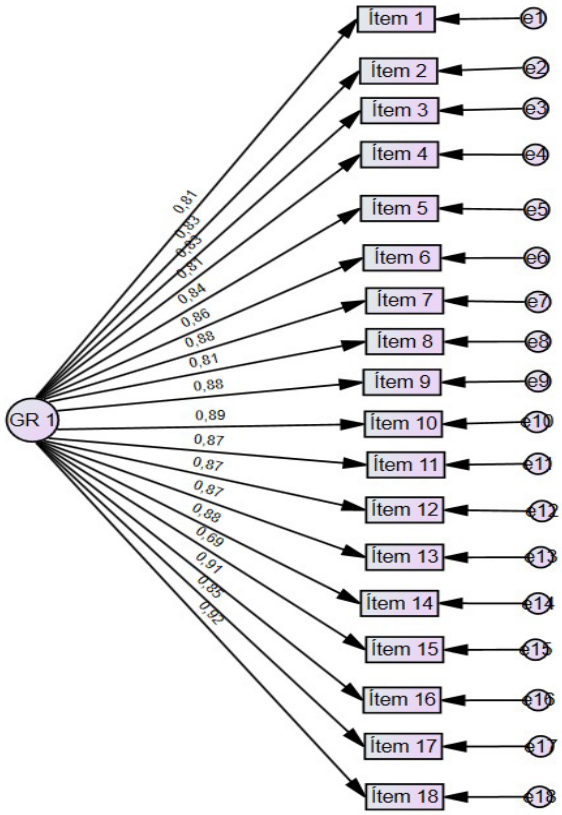
La validez del constructo de la prueba la hemos obtenido a través de la realización de un análisis factorial exploratorio (tabla 4). El modelo teórico extraído en el análisis factorial exploratorio (AFE) ha sido contrastado mediante la realización de un análisis factorial confirmatorio (AFC). En la figura 1 se representa el diagrama estructural planteado, en el que se puede observar los índices de correlación existente entre ítem y dimensión.

TABLA 4. Matriz de componente rotado

Ítems	Valor Dimensión
1	.898
2	.910
3	.873
4	.861
5	.917
6	.917
7	.923
8	.899
9	.914
10	.940
11	.924
12	.930
13	.927
14	.922
15	.846
16	.932
17	.908
18	.944

Fuente: elaboración propia

FIGURA 1. Diagrama estructural del Cuestionario de Opinión del Alumnado sobre el Grado de Actuación Docente del Profesorado



Fuente: elaboración propia

Las cargas factoriales de las dimensiones se sitúan entre .69 y .92, resultados que demuestran que existe un alto grado de correlación entre cada uno de los ítems y la única dimensión de la que se compone el cuestionario.

La tabla 5 muestra los valores obtenidos y de referencia para el ajuste del modelo según Levy Mangin et al. (2006): chi-cuadrado (CMIN), Índice de bondad de ajuste (GFI), índice de bondad de ajuste parsimónico (PGFI), índice de ajuste normalizado (NFI) e índice de ajuste parsimónico normalizado (PNFI).

TABLA 5. Índices de ajuste

Índice	Resultado	Ajuste
CMIN		CMIN<500
GFI		GFI>.7
PGFI		PGFI>.7
NFI		NFI>.7
PNFI		PNFI>.7

Fuente: elaboración propia

4. DISCUSIÓN Y CONCLUSIONES

En este estudio se ha analizado cuál es la validez confirmatoria del Cuestionario de Opinión del Alumnado sobre la Actuación Docente del Profesorado.

En este sentido, cabría destacar, tal como apunta Arranz Val, (2007) que la sociedad reclama con mayor insistencia que las universidades funcionen en base a una serie de criterios, los cuales se encuentren vinculados con la eficacia, la eficiencia y la excelencia. Con lo que el objetivo fundamental de la investigación estuvo centrado en validar el instrumento diseñado por la Unidad General de Calidad de la Universidad de Sevilla, para verificar que permite evaluar la intervención del docente universitario en cada una de sus asignaturas. Para ello, y teniendo en cuenta la finalidad del instrumento, parece importante apoyarnos en procesos de evaluación de las metodologías que den respuesta, por tanto, a las necesidades educativas y profesionales, generadas por la implantación del actual EEES.

La validación de instrumento de medida de la calidad docente del profesorado permitirá a las instituciones educativas ser un elemento clave para optimizar el funcionamiento de la actuación docente del profesorado, así como de su propia intervención, permitiendo así poder llevar a cabo una valoración objetiva de los resultados (García Ramos y Congosto Luna, 2000).

Todo ello resulta de vital importancia si tenemos en cuenta que la evaluación de la enseñanza realizada por los estudiantes (Students´evaluations of teaching SET) es utilizada, en la mayoría de las universidades de todo el mundo, como una medida de calidad docente, de manera que los resultados de las evaluaciones de las valoraciones del alumnado sobre la actuación docente del profesorado ofrecen un feedback al docente para mejorar su enseñanza (Balam y Shannon, 2010).

La validación del cuestionario nos permitirá contar con un instrumento que dará la posibilidad de valorar la intervención docente del profesorado universitario, de forma que los resultados obtenidos en las evaluaciones realizadas por el alumnado podrían ser de gran utilidad para los docentes, proporcionando datos objetivos. Por ello, resulta de vital importancia que se tengan en cuenta tanto los aspectos relacionados con el proceso de aplicación del instrumento, como el uso que a posteriori se hace de los resultados, resultando necesario que tras su aplicación haya un proceso de retroalimentación hacia las partes implicadas (Hattie, 2008; Leung et al., 2013; Muñiz y Fonseca-Pedrero, 2008). Por ello, se entiende siempre este como una estrategia que permita mejorar la calidad de la enseñanza, implicando en ella a los estudiantes y permitiendo así orientar la actividad docente (Padilla Carmona y Gil Flores, 2008). Todo optimizando la calidad docente universitaria y pudiendo así alcanzar una evaluación más formativa (Moore et al., 2012).

Tras los diferentes análisis realizados, se ha estimado que el cuestionario presenta unas adecuadas propiedades psicométricas que le dan validez para la valoración de la actuación docente universitaria, se obtiene una buena consistencia interna, con un valor mayor a .70 y menor o igual a .95 (.988), lo que hace que garantiza el alto índice de fiabilidad del instrumento, comprobando la validez del modelo y comprobando la estructura confirmatoria de este.

Por este motivo, la robustez científica del instrumento lo convierte en idóneo para:

- Establecer una herramienta estandarizada que permita recoger información sobre la opinión que tiene el alumnado sobre el grado de satisfacción docente del profesorado.
- Establecer una base sólida, basada en una serie de evidencias, que permita guiar al docente y a la propia institución educativa sobre el grado de calidad de la acción docente llevada a cabo con el alumnado.
- Generar un modelo evaluativo común en este tipo de temas, unificando así diferentes y modelos utilizados hasta el momento, por instituciones educativas de Andalucía y España, donde los estándares de evaluación se agrupan bajo otras dimensiones o parámetros.
- Crear un punto de partida que permita poner de manifiesto la importancia de la calidad docente universitaria, en las aulas, en nuestros días y cómo esta afecta al desarrollo de la acción motivacional del alumnado.

Todo ello debe ser reforzado con aspectos donde, además del criterio del alumnado, se tengan en cuenta otro tipo de valoraciones, como son: las autoevaluaciones del propio profesorado, informes de evaluadores externos a la institución, compañeros de trabajo, etc., consiguiendo así que la evaluación docente llevada a cabo se complemente y tenga un valor más integral. Es así como se podrá llevar a cabo una evaluación más rica y valiosa que facilitará, a la propia institución educativa en general, y al docente en particular, reorientar, en caso de necesidad, su acción educativa a través de la propuesta de planes de mejora.

Por último, hay que apuntar que la principal limitación del cuestionario radica en la limitación de la muestra, y en que la misma ha sido reclutada en un mismo centro universitario, aunque en distintas titulaciones de Grado y Máster. Por ello, para futuras investigaciones, puede resultar interesante replicar el estudio en otras titulaciones pertenecientes a otras universidades, a nivel autonómico y nacional. Asimismo, resaltar que es significativo que, en los tiempos actuales, no se concibe en el actual contexto que la herramienta analizada que evalúa al profesorado no advierta de una dimensión o concepto de análisis que contemple la competencia digital docente.

5. REFERENCIAS

Agencia Nacional de Evaluación de la Calidad y Acreditación. (2024). *Programa ACADEMIA: Procedimiento de acreditación para el acceso a los cuerpos docentes universitarios*. https://www.aneca.es/procedimiento-de-acreditación-2024

Ajjawi, R., & Boud, D. (2017). Researching feedback dialogue: an interactional analysis approach. *Assessment & Evaluation in Higher Education, 42*(2), 252-265. https://doi.org/10.1080/02602938.2015.1102863

Aláez, M., García-Olalla, A., García-Feijoo, M., y Eizaguirre, A. (2018). Evaluación de la calidad docente en una facultad de derecho. *Revista de Educación y Derecho, 17*, 1-28. https://doi.org/10.1344/REYD2018.17.23482

Álvarez, V., García, E., y Gil, J. (1999). Características de la docencia mejor evaluada por los alumnos en las diferentes áreas de enseñanza universitaria. *Revista Española de Pedagogía, 214,* 445-464.

Arranz Val, P. (2007*). Los sistemas de garantía de calidad en la educación superior en España. Propuesta de un modelo de acreditación para las titulaciones de grado en empresa*. Servicio de Publicaciones, Universidad de Burgos. https://doi.org/10.36443/10259/78

Balam, E. M., & Shannon, D. M. (2010). Student ratings of college teaching: A comparison of faculty and their students. *Assessment & Evaluation in Higher Education, 35*(2), 209-221. https://doi.org/10.1080/02602930902795901

Campos Soto, N., Ramos Navas-Parejo, M., y Moreno Guerrero, A.J. (2020). Realidad virtual y motivación en el contexto educativo: Estudio bibliométrico de los últimos veinte años de Scopus. *Alteridad, 15*(1), 47-60. https://doi.org/10.17163/alt.v15n1.2020.04

Carmines, E., y Zeller, R. (1979). *Reliability and validity assessmen*t. Thousand Oaks. SAGE Publications.

Cervantes Mata, C.M. (2020). El nuevo rol del profesor universitario en la etapa postcovid: formación y desarrollo de competencias docentes. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo, 14*(28), e617. https://doi.org/10.23913/ride.v14i28.1804

Clayson, D. E. (2009). Student Evaluations of Teaching: Are They Related to What Students Learn?: A Meta-Analysis and Review of Literature. *Journal of Marketing Education, 31*(1), 16-30. https://doi.org/10.1177/0273475308324086

De-Juanas Oliva, Á., y Beltrán Llera, J. A. (2013). Valoraciones de los estudiantes de ciencias de la educación sobre la calidad de la docencia universitaria. *Educación XX1, 17*(1), 59-82. https://doi.org/10.5944/educxx1.17.1.10705

Delgado-López-Cózar, E. (2022). Ante la reforma del sistema español de evaluación académica: recordando el pasado para no errar en el presente. *Anuario ThinkEPI, 16*. https://doi.org/10.3145/thinkepi.2022.e16a37

Dorta González M. I., y Dorta González, P. (2012). Evaluación de la actividad docente universitaria considerando la percepción que los estudiantes tienen de sus compañeros. *Revista Iberoamericana de Evaluación Educativa, 5*(2), 147-159. https://doi.org/10.15366/riee2012.5.2.010

Elitok, A., & Cavus, B. (2019). University Life Quality and Impact Areas. *Universal Journal of Educational Research,* 7(6), 1376-1386. https://doi.org/10.13189/ujer.2019.070605

García Berro, E., Colom Pastor, X., Martínez Solanas, É., Sallarés, J., y Roca Martín, S. (2011). La encuesta al alumnado en la evaluación de la actividad docente del profesorado. *Aula abierta, 39*(3), 3-14.

García-Herrero, M., Rodríguez-Conde, M.J., y Martínez-Abad, F. (2024). Factores de calidad docente asociados a la equidad educativa: formación del profesorado y estrategias docentes. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 27*(1), 75-88. https://doi.org/10.6018/reifop.595181

García Ramos, J. M., y Congosto Luna, E. (2000). Evaluación y calidad del profesorado. En T. González Ramírez (Coord.). *Evaluación y gestión de la calidad educativa: un enfoque metodológico* (pp. 127-158). Aljibe.

García Garduño, J. y Medécigo Shej, A. (2014). Los criterios que emplean los estudiantes universitarios para evaluar la ineficacia docente de sus profesores. *Perfiles Educativos, 36*(143), 124-139. https://doi.org/10.1016/S0185-2698(14)70613-0

González López, I., y López Cobo, I. (2010). Validación y propuesta de un modelo de indicadores de evaluación de la calidad en la universidad. *Revista Iberoamericana de Educación*, 53(6), 1-13. https://doi.org/10.35362/rie5361706

Hair, J., Black, W., Babin, B., y Anderson, R. (2010). *Multivariate data analysis*. Prentice-Hall.

Harvey, L. (2003). Student feedback. *Quality in Higher Education, 9* (1), 3-20. https://doi.org/10.1080/13538320308164

Hattie, J. (2008). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge. https://doi.org/10.4324/9780203887332

Hernández-Sampieri, R., Fernández-Collado, C., y Baptista-Lucio, P. (2014). *Metodología de la Investigación*. McGraw-Hill.

Jiménez, B., Bordas, I., Coronel, J.M., Domínguez, G., Gairin, J.; Pio, A., Santos, M.A., Tejada, J. (1999). *Evaluación de programas, centros y profesores*. Síntesis.

Leung, X. Y., Jiang, L., & Busser, J. (2013). Online student evaluations of hospitality professors: A cross-cultural comparison. *Journal of Hospitality, Leisure, Sport & Tourism Education, 12*(1), 36-46. https://doi.org/10.1016/j.jhlste.2012.10.001

Levy, J.P., y Varela, J. (2003). *Análisis multivariante para las ciencias sociales*. Pearson.

Lévy Mangin, J.-P., Varela Mallou, J., Abad González, J., y Ábalo Piñeiro, J. (2006). *Modelización con estructuras de covarianzas en Ciencias Sociales: Temas esenciales, avanzados y aportaciones especiales*. Netbiblo.

López Cámara, A. B., González-López, I., y León-Huertas, C. de. (2015). Un análisis factorial exploratorio para la construcción de un modelo de indicadores de evaluación docente universitaria. *Culture and Education, 27*(2), 351-364. https://doi.org/10.1080/11356405.2015.1035546

Mat Daud, N., y Abu Kassim, N. L. (2011). Examining student rating of teaching effectiveness using FACETS. *Journal of Applied Measurement, 12*(2), 135-143.

Marqués, R. (2008). *Profesores/as motivados/as. Un liderazgo positivo promueve el bienestar docente*. Narcea.

Mateo, J. (2000). La evaluación del profesorado y la gestión de la calidad de la educación. Hacia un modelo comprensivo de evaluación sistemática de la docencia. *Revista de Investigación Educativa*, *18*(1), 7–34. https://revistas.um.es/rie/article/view/121451

Ministerio de Universidades. (2023*).* Real Decreto 678/2023, de 18 de julio, por el que se regula la acreditación estatal para el acceso a los cuerpos docentes universitarios y el régimen de los concursos de acceso a plazas de dichos cuerpos. *Boletín Oficial del Estado, 213,* de 6 de septiembre de 2023. https://www.boe.es/eli/es/rd/2023/07/18/678

Molero López-Barajas, D., y Ruiz Carrascosa, J. (2005). La evaluación de la docencia universitaria. Dimensiones y variables más relevantes. Revista de Investigación Educativa, 23(1), 57-84.

Montero, L., Triviño, X., Dois, A., Sirhan, M., y Leiva, L. (2017). Percepción de los académicos del rol docente del médico*. Investigación en Educación Médica, 6*(23), 198-205. https://doi.org/10.1016/j.riem.2017.01.001

Moore, S., Walsh, G., y Rísquez, A. (2012). *Estrategias eficaces para enseñar en la Universidad. Guía para docentes comprometidos*. Narcea.

Muñiz, J., y Fonseca-Pedrero, E. (2008). Construcción de instrumentos de medida para la evaluación universitaria. *Revista de investigación en educación, 5*(1), 13-25.

Muñoz Cantero, J. M., Ríos de Deus, M. P., y Abalde, E. (2002). Evaluación Docente vs. Evaluación de la Calidad. *Revista Electrónica de Investigación y Evaluación Educativa, 8*(2), 103-134.

Pozo Muñoz, C., Giménez Torres, M. L., y Bretones Nieto, B. (2009). *La evaluación de la calidad docente en el nuevo marco del EEES. Un estudio sobre la encuesta de opinión del Programa DOCENTIA-ANDALUCÍA*. http://hdl.handle.net/10272/4905

Navarro C., y Ramírez, M.S. (2018). Mapeo sistemático de la literatura sobre evaluación docente (2013-2017). *Educaçao e Pesquisa*, *44*, e185677. https://doi.org/10.1590/S1678-4634201844185677

O'Dwyer, L., y Bernauer, J. (2014). Quantitative research for the qualitative researcher. Sage publications. https://doi.org/10.4135/9781506335674

Padilla Carmona, M. T., y Gil Flores, J. (2008). La evaluación orientada al aprendizaje en la Educación Superior: condiciones y estrategias para su aplicación a la Docencia Universitaria. *Revista Española de Pedagogía, 66*(241), 467-485.

Pozo Muñoz, C., Bretones Nieto, B., Martos Méndez, M. J., y Alonso Morillejo, E. (2011). Evaluación de la actividad docente en el Espacio Europeo de Educación Superior: un estudio comparativo de indicadores de calidad en universidades europeas. *Revista Española de Pedagogía, 69*(248), 145-163.

Rivero Castro, M.R., y Medeiros, M.P. (2023). ¿Calidad en la Formación Inicial Docente?: Análisis de los nuevos Estándares de la Profesión Docente para carreras de Pedagogía. *Perspectiva Educacional, 62*(3), 3-26. http://dx.doi.org/10.4151/07189729-vol.62-iss.3-art.1248

Rodríguez Torres, A.F., Rodríguez Alvear, J.C., Arias Moreno, E.R. (2020). El profesorado universitario novel en shock: propuesta de mejora revista sistemática. *Revista Educare, 24*(1), 245-269. https://doi.org/10.46498/reduipb.v24i1.1245

Ruiz Díaz, M. A., Pardo Merino, A., y San Martín Castellanos, R. (2010). Modelos de ecuaciones estructurales*. Papeles del Psicólogo, 31*(1), 34-45.

Tejedor, F. J., y García-Valcárcel, A. (1996). La evaluación de la calidad de la docencia universitaria en el marco de la evaluación institucional desde la perspectiva del alumno. En F.J. Tejedor y J.L. Rodríguez Diéguez (Coords.), *Evaluación educativa II: Evaluación institucional fundamentos teóricos y aplicaciones prácticas* (pp. 93-122). IUCE.

Universidad de La Rioja, (2024). *Encuestas de satisfacción de los estudiantes con la actividad docente del profesorado.* Oficina de Calidad y Evaluación. https://www.unirioja.es/adminis-

tracion-y-servicios/oficina-de-calidad-y-evaluacion/evaluacion-y-calidad-del-profesorado/

Urkidi, P., Losada, D., López, V., y Yuste, R. (2020). El Acceso a la Formación Inicial del Profesorado y la Mejora de la Calidad Docente. *Revista Complutense de Educación, 31* (3), 353-364. https://doi.org/10.5209/rced.63476

Walton-Fisette, J. L., & Sutherland, S. (2020). Time to Shape Up: Developing Policies, Standards and Practices That Are Socially Just. *Physical Education and Sport Pedagogy, 25*(3), 274-287. https://doi.org/10.1080/17408989.2020.1741531

6. ANEXO

Código: H5-E1	EVALUCIÓN Y MEJORA DE LA CALIDAD DE LA ENSEÑANZA Y EL PROFESORADO: CUESTIONARIO DE OPINIÓN DE LOS ESTUDIANTES SOBRE LA ACTUACIÓN DOCENTE DEL PROFESORADO	Versión: 1 Fecha:

Cuestionario de opinión del alumnado sobre la actuación docente del profesorado

Profesor/a: ……

Asignatura: ……

Titulación: ……

Sexo:		Hombre		Mujer

Edad	Decena	1	2	3	4	5	6	7	8	9	
	Unidad	1	2	3	4	5	6	7	8	9	0

Curso más alto en el que está matriculado	1º	2º	3º	4º	5º	6º
Curso más bajo en el que está matriculado	1º	2º	3º	4º	5º	6º

La asignatura me interesa:	Nada	Algo	Bastante	Mucho

El grado de dificultad de esta asignatura es:	Bajo	Medio	Alto	Muy alto

Asisto a clase:	Menos del 25%	26-50%	51-75%	Más del 75%

A la Atención de Alumnos de este profesor he asistido:	Ninguna vez	Una vez	2-3 veces	Más de 3 veces

A continuación, le presentamos una serie de cuestiones relativas a la docencia de este/a profesor/a en esta asignatura. Su colaboración es necesaria y consiste en señalar en la escala de respuesta su grado de acuerdo con cada una de las afirmaciones, teniendo en cuenta que 1 significa el menor grado de acuerdo y 5 significa el mayor grado de acuerdo. Si no tiene suficiente información sobre el enunciado o considera que no procede, marque la opción NS.

1. Me ha dado orientaciones para conocer el proyecto docente de la asignatura	1	2	3	4	5	NS
2. Su docencia se ajusta a la planificación prevista en el proyecto docente	1	2	3	4	5	NS
3. Me atiende adecuadamente en la hora de Atención a Alumnos	1	2	3	4	5	NS
4. Cumple adecuadamente el horario de Atención a Alumnos	1	2	3	4	5	NS
5. La bibliografía y demás material docente recomendado me están resultando útiles para el seguimiento de la asignatura	1	2	3	4	5	NS
6. Su docencia está bien organizada	1	2	3	4	5	NS
7. Los medios que utiliza para impartir su docencia son adecuados	1	2	3	4	5	NS
8. La bibliografía y demás material docente recomendado están a disposición de los estudiantes	1	2	3	4	5	NS
9. Explica con claridad	1	2	3	4	5	NS
10. Se interesa por el grado de comprensión de sus explicaciones	1	2	3	4	5	NS
11. Expone ejemplos para poner en práctica los contenidos de la asignatura	1	2	3	4	5	NS
12. Resuelve las dudas que se le plantean	1	2	3	4	5	NS
13. Fomenta un clima de trabajo y participación	1	2	3	4	5	NS
14. Motiva a los/las estudiantes para que se interesen por la asignatura	1	2	3	4	5	NS
15. Trata con respeto a los/las estudiantes	1	2	3	4	5	NS
16. Su docencia me está ayudando a alcanzar los objetivos de la asignatura	1	2	3	4	5	NS
17. Los criterios y sistemas de evaluación me parecen adecuados	1	2	3	4	5	NS
18. En general, estoy satisfecho/a con la actuación docente desarrollada por este/a profesor/a	1	2	3	4	5	NS