Generative artificial intelligence in higher education learning: A review based on academic databases

Daniel Andrade-Girón¹, William Marín-Rodriguez¹, *, Juana Sandivar-Rosas², Edgardo Carreño-Cisneros¹, Edgar Susanibar-Ramirez¹, Marcelo Zuñiga-Rojas¹, Julio Angeles-Morales³, Henry Villarreal-Torres³

¹ Universidad Nacional José Faustino Sánchez Carrión, Huacho. Lima, Perú.
² Universidad Nacional Mayor de San Marcos. Lima, Perú.
³ Universidad San Pedro, Chimbote. Ancash, Perú.
*Corresponding author: wmarin@unjfsc.edu.pe (Email), https://orcid.org/0000-0002-0861-9663 (ORCID).

ABSTRACT

Objective. The rapid integration of Generative Artificial Intelligence (AI), especially tools like ChatGPT, into educational sectors has spurred significant academic interest. This review article systematically examines the current scholarly landscape concerning the use of ChatGPT within higher education.

Design/Methodology/Approach. Drawing from various academic databases between 2022 and 2024, we meticulously adhere to PRISMA guidelines, evaluating a final set of 28 out of 1740 initial articles based on predetermined inclusion and exclusion criteria.

Results/Discussion. Our analysis reveals diverse global contributions predominantly from Asia and identifies a prevalent quantitative research approach among the studies. We delve into the selected articles’ geographical distribution, methodologies, and thematic outcomes, highlighting a notable lack of research from Latin America. The review critically assesses the validity, utility, and time optimization aspects of ChatGPT in educational settings, uncovering a positive impact on student learning and time management. However, we pinpoint a significant gap in rigorous experimental research, underscoring the need for studies with random sampling and controlled settings to enhance the external validity of findings. Additionally, we call attention to the ethical considerations and the necessity for higher education institutions to adapt teaching methodologies to incorporate AI effectively.

Conclusion. The article concludes with recommendations for future research to address the identified gaps and optimize the educational use of generative AI technologies like ChatGPT.

Keywords: generative artificial intelligence, higher education, ChatGPT, educational technology, academic databases

Received: 21-11-2023, Accepted: 21-03-2024, Published: 05-04-2024

Editor: Adilson Luiz Pinto


Copyright: © 2024 The author(s). This is an open access article distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming, and building upon the material as long as the license terms are followed.
INTRODUCTION

The history of artificial intelligence has its roots in the 1950s (Grzybowski, Pawlikowska–Łagód, & Clark, 2024), with significant contributions from notable figures such as Alan Turing and John McCarthy (Meadows & Sternfeld, 2023). Turing proposed revolutionary ideas in 1950 by introducing a test designed to evaluate the ability of a machine to exhibit intelligent behavior similar to that of a human being, called the Turing test (Gonçalves, 2023), while McCarthy, in 1956, not only coined the term “artificial intelligence” (AI), but also defined the field as the science aimed at creating intelligent machines, especially computer programs capable of emulating human thought and behavior (Mohammed et al., 2024).

For more than 67 years, artificial intelligence research has made notable achievements in theory and practical, real-world applications (Jiang et al., 2022). AI has been integrated into many activities, and its management is becoming essential in organizations (Vasquez, 2022; Linden, Tilman, & Laurent, 2023; Auza-Santiváñez et al., 2023). This is reflected in the growth of the global AI market, valued at $150.2 billion in 2023 and projected to increase at a CAGR of 36.8% between 2023 and 2030 (Dou et al., 2023).

AI plays a crucial role in driving the advancement of science and technology (Lu, 2019; Gruetzemacher & Whittlestone, 2022), which has a significant impact on multiple industries (Chen et al., 2024), positioning itself as a critical driver for emerging technologies such as big data analytics, robotics and the internet of things (IoT) (Özdemir & Hekim, 2018). In addition to the rise of generative AI tools such as ChatGPT (Gomez Cano et al., 2023; Polyporit & Pahos, 2024), consequently the significant impact on the science of education and society (Jaiswal & Arun, 2021; Crompton & Song, 2021; Kumar et al., 2023; Junco Luna, 2023), this panorama raises questions about the use of AI tools in the teaching-learning process (Haque et al., 2023; Wen, 2024; Shamsuddinova, Heryani, & Naval, 2024).

Generative AI poses new challenges for teachers in the teaching and research process (Hwang et al., 2020). The advancement of intelligent agents, such as robust text generation systems (Yu et al., 2022), systems capable of generating coherent and contextually appropriate responses from user questions and comments are used in various educational applications (Ray, 2023). This highlights the existing gap in current educational models and the need for a new type of professional with skills to handle AI technologies in information management, orienting towards an approach focused on knowledge management (Li & Gu, 2023; González-Valiente, 2023, Panduro, 2023).

The growing research publication on applying generative AI, such as ChatGPT, in education highlights the importance of conducting systematic reviews and meta-analyses (Thorpe, 2023). Although several studies have been conducted on using ChatGPT in various educational settings (Bin-Nashwan, Sadallah, & Bouteraa, 2023; Bouker, 2024), their scope has been limited. It has not provided a comprehensive overview of this technology’s possible benefits and limitations in these fields (Aithal & Aithal, 2024). Furthermore, policy implementers in education, such as ChatGPT users’ opinions, are divided regarding adopting this technology in education (Oliva et al., 2022; Fuchs, 2023; Rudolph, Tan, & Tan, 2023; Larrosa et al., 2023). The lack of consensus on best practices for its implementation in higher education and the need to address the ethical implications of its use in educational practice has not yet been resolved (Rane et al., 2023; Vieytes, 2023).

Several investigations have been carried out that focus on systematic reviews of the topic. For example, Perera and Lankathilaka (2023) examined both benefits and drawbacks. On the other hand, Imran and Almusharraf (2023) analyzed the related opportunities and challenges. Additionally, Vargas-Murillo, de la Asunción, and de Jesús Guevara-Soto (2023) investigated the topic’s impact, benefits, and use. However, none of the analyzed articles have addressed crucial aspects such as validity, usefulness, and time optimization in applying ChatGPT in higher education.

Therefore, using academic databases as data sources, a systematic review of the use of ChatGPT in higher education is warranted to identify knowledge gaps and guide future research in this area. This review aims to identify the acceptance, validity, usefulness, and
time optimization of ChatGPT in higher education presented in the scientific literature.

**METHODOLOGY**

This research has been carried out following the systematic review methodology proposed in the literature (Pigott & Polanin, 2020; Sánchez, 2010). This methodology has been based on the guidelines established by PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), as documented by Serrano, Navarro, & González (2022); Schwarzer, Carpenter, & Rücker (2015), and Alexander (2020).

In this context, an exhaustive literature review was carried out to analyze the most recent publications between 2022 and 2024. To conduct this review systematically, meticulous planning was required, following the guidelines outlined by Brereton et al. (2009).

A crucial step in this process has been the precise formulation of the research objective since clarity in the research questions, and their components is essential for a successful systematic review. A detailed research protocol has been developed that has comprehensively established the design of the systematic review. This protocol has rigorously addressed the following aspects: the study selection criteria, the sources of information used in the bibliographic search, the research strategies implemented, and the procedures for collecting and analyzing the data obtained.

An exhaustive search was conducted in specialized databases to locate relevant information supporting our research (See Table 1). Table 2 presents the implemented search strategy in detail.

In this scientific research, inclusion and exclusion criteria refer to the predefined patterns and guidelines used to discern which studies or articles will be incorporated into the systematic review and which will be excluded (See Table 3).

After meticulously applying the inclusion and exclusion criteria, a rigorous restriction was carried out on the sample to analyze only those articles that provided relevant and consistent information with the purpose of the research. The initial process, as detailed in the accompanying flowchart (Figure 1), revealed the presence of 1,740 articles in the six databases examined. Subsequently, by eliminating

<table>
<thead>
<tr>
<th>IDE</th>
<th>Database</th>
<th>No. documents</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB1</td>
<td>Scopus</td>
<td>214</td>
<td>12.4</td>
</tr>
<tr>
<td>DB2</td>
<td>IEEE Xplore Digital Library</td>
<td>51</td>
<td>2.9</td>
</tr>
<tr>
<td>DB3</td>
<td>ScienceDirect</td>
<td>818</td>
<td>47.0</td>
</tr>
<tr>
<td>DB4</td>
<td>Wiley</td>
<td>401</td>
<td>23.0</td>
</tr>
<tr>
<td>DB5</td>
<td>Pubmed</td>
<td>35</td>
<td>2.0</td>
</tr>
<tr>
<td>DB6</td>
<td>Sage Journals</td>
<td>221</td>
<td>12.7</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>1,740</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 1. Database consulted.**

<table>
<thead>
<tr>
<th>Database</th>
<th>Search query</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scopus</td>
<td>TITLE-ABS-KEY ( chatgpt AND in AND higher AND education AND learning )</td>
</tr>
<tr>
<td>IEEE Xplore Digital Library</td>
<td>(“All Metadata”:ChatGPT) AND (“All Metadata”:Higher) AND (“All Metadata”:Education) AND (“All Metadata”:Learning)</td>
</tr>
<tr>
<td>ScienceDirect</td>
<td>ChatGPT AND Higher AND Education AND Learning</td>
</tr>
<tr>
<td>PubMed</td>
<td>(((ChatGPT) AND (Higher)) AND (Education)) AND (Learning)</td>
</tr>
<tr>
<td>Wiley</td>
<td>“ChatGPT” anywhere and “Higher” anywhere and “Education” anywhere and “Learning” anywhere</td>
</tr>
<tr>
<td>IOPSience</td>
<td>ChatGPT AND Higher AND Education AND Learning</td>
</tr>
<tr>
<td>Sage Journals</td>
<td>ChatGPT AND Higher AND Education AND Learning</td>
</tr>
</tbody>
</table>

**Table 2. Search formula for each database.**
Characteristics | Inclusion | Exclusion
---|---|---
Participants | Higher education students | Non-educational institutions of higher education
Phenomenon of interest | Usage of ChatGPT in higher education | Use of systems other than ChatGPT in higher education
Period | Studies: from 2022 to 2024 | Studies outside this time range
Language | English | Non-English languages

Table 3. Inclusion and exclusion criteria.

![Flowchart of the search and reference selection method of the systematic review.](image)

Number of additional records identified through other sources
(n = 0)

Number of records after deleting duplicates
(n = 1450)

Number of records screened
(n = 1450)

Number of full-text articles excluded:
Not fit for purpose (n = 76)

Number of full text articles to evaluate your eligibility
(n = 104)

Articles included in the systematic review
(n = 28)

Results and discussion
Table 4 shows the most notable characteristics in the context of the systematic review carried out. This analysis covers a variety of fundamental attributes, including authorship, year of
<table>
<thead>
<tr>
<th>Author &amp; Year</th>
<th>Country</th>
<th>Research level type</th>
<th>Sample</th>
<th>Sample selection</th>
<th>Control group</th>
<th>Experiment group</th>
<th>Reliability and Validity</th>
<th>Test - Test</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Kayalı, Yavuz, Balat, &amp; Çalışan, 2023)</td>
<td>Turkey</td>
<td>Mix Multimethod</td>
<td>84</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Students had positive experiences using ChatGPT in education.</td>
</tr>
<tr>
<td>(Yilmaz &amp; Yilmaz, 2023)</td>
<td>Turkey</td>
<td>Quantitative Experimental: pretest-postest</td>
<td>45</td>
<td>Random</td>
<td>21</td>
<td>24</td>
<td>Not specific</td>
<td>ANOVA-ANCOVA</td>
<td>ChatGPT in education was beneficial to students’ learning process and outcomes.</td>
</tr>
<tr>
<td>(Habibi, Muhaimin, Danibao, &amp; Wibowo, 2023)</td>
<td>Indonesia</td>
<td>Quantitative</td>
<td>1117</td>
<td>Simple random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Structural equations model</td>
<td>Meaningful relationships for using ChatGPT in learning.</td>
</tr>
<tr>
<td>(Chiu F., 2024)</td>
<td>Japan</td>
<td>Quantitative: Thematic analysis</td>
<td>51</td>
<td>Not specific</td>
<td>No</td>
<td>No</td>
<td>Not specific</td>
<td>Not specific</td>
<td>GenAI apps can quickly finish learning activities.</td>
</tr>
<tr>
<td>(Niloy, y otros, 2024)</td>
<td>Bangladesh</td>
<td>Mix Triangulation</td>
<td>422</td>
<td>Not random, stratified</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>SEM</td>
<td>Demonstrates strong positive associations with students’ intention to use ChatGPT.</td>
</tr>
<tr>
<td>(Kumar, Rao, Singhania, Verma, &amp; Kheterpal, 2024)</td>
<td>India</td>
<td>Mixto</td>
<td>45</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Triangulation</td>
<td>ChatGPT improves pedagogical innovation, and academic integrity.</td>
</tr>
<tr>
<td>(Bouker, 2024)</td>
<td>Morocco</td>
<td>Quantitative</td>
<td>319</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Partial least squares</td>
<td>The perceived usefulness of ChatGPT positively influences student satisfaction.</td>
</tr>
<tr>
<td>(Habib, Vogel, Anili, &amp; Thorne, 2024)</td>
<td>USA</td>
<td>Mix</td>
<td>100</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>t-student</td>
<td>AI helps with divergent thinking, an important part of the creative process.</td>
</tr>
<tr>
<td>(Remoto, 2024)</td>
<td>Philippines</td>
<td>Quantitative</td>
<td>15</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>The results had promising implications for education.</td>
</tr>
<tr>
<td>(Grájeda, Burgos, Córdova, &amp; Sanjinés, 2024)</td>
<td>Bolivia</td>
<td>Quantitative</td>
<td>4127</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Confirmatory factor analysis</td>
<td>Results indicate that AI tools have a significant impact.</td>
</tr>
</tbody>
</table>

Table 4. Descriptive characteristics of the included studies.

(Continued)
<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Research level type</th>
<th>Sample</th>
<th>Sample selection</th>
<th>Control group</th>
<th>Experiment group</th>
<th>Reliability and Validity</th>
<th>Test -Test</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michel-Villarreal, Vilalta-Perdomo, Salinas-Navarro, Thierry-Aguilera, &amp; Gerardou, 2023</td>
<td>UK</td>
<td>Quantitative</td>
<td>No</td>
<td>Not specific</td>
<td>No</td>
<td>No</td>
<td>Not specific</td>
<td>Etnographic</td>
<td>The findings of this study highlighted the transformative potential of ChatGPT in education.</td>
</tr>
<tr>
<td>Singh, Tayarani-Najaran, &amp; Yaqoob, 2023</td>
<td>UK</td>
<td>Quantitative: Relational</td>
<td>430</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Correlation</td>
<td>If used correctly, it can have many positive impacts, and if used poorly, it can harm students.</td>
</tr>
<tr>
<td>Kiryakova &amp; Angelova, 2023</td>
<td>Bulgaria</td>
<td>Quantitative: Relational</td>
<td>87</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Monte Carlo</td>
<td>Los resultados del estudio actual muestran que ChatGPT tiene el potencial de apoyar la enseñanza y el aprendizaje.</td>
</tr>
<tr>
<td>Hasanein &amp; Sobaih, 2023</td>
<td>Saudi Arabia</td>
<td>Quantitative</td>
<td>85</td>
<td>Not random: intentional</td>
<td>No</td>
<td>No</td>
<td>Not specific</td>
<td>Topic analysis</td>
<td>ChatGPT serves as an adaptable resource for both students and teachers.</td>
</tr>
<tr>
<td>Xu, Wang, Zhang, Zhang, &amp; Wu, 2023</td>
<td>China</td>
<td>Quantitative</td>
<td>8</td>
<td>Not random: intentional</td>
<td>No</td>
<td>No</td>
<td>Not specific</td>
<td>Topic analysis</td>
<td>ChatGPT can assist in pedagogical adjustment to align PLEs with formal education.</td>
</tr>
<tr>
<td>Kelly, Sullivan, &amp; Strampel, 2023</td>
<td>Australia</td>
<td>Quantitative</td>
<td>1135</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>U de Mann-Whitney</td>
<td>GenerAI tools have significantly altered teaching and learning practices.</td>
</tr>
<tr>
<td>Hmoud, Swaity, Hamad, Karram, &amp; Daher, 2024</td>
<td>Palestine</td>
<td>Quantitative</td>
<td>15</td>
<td>Not random</td>
<td>No</td>
<td>No</td>
<td>Not specific</td>
<td>MAXQDA 2022</td>
<td>The research results revealed that the implementation of ChatGPT had a positive impact.</td>
</tr>
<tr>
<td>Valova, Mladenova, &amp; Kavev, 2024</td>
<td>Bulgaria</td>
<td>Quantitative</td>
<td>102</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>The integration of AI technologies in education concludes.</td>
</tr>
<tr>
<td>Bower, Torrington, Lai, Petocz, &amp; Alfano, 2024</td>
<td>Australia</td>
<td>Mix</td>
<td>318</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Kappa de Cohen</td>
<td>Generative AI significantly influenced teaching and assessment.</td>
</tr>
</tbody>
</table>

Table 4. (Continued)
<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Research level type</th>
<th>Sample</th>
<th>Sample selection</th>
<th>Control group</th>
<th>Experiment group</th>
<th>Reliability and Validity</th>
<th>Test - Test</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Chan &amp; Lee, 2023)</td>
<td>Hong Kong</td>
<td>Mix</td>
<td>583</td>
<td>Not random: for convenience</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>T de student</td>
<td>The study highlighted the importance of combining technology with teaching methods.</td>
</tr>
<tr>
<td>(Wang, y otros, 2023)</td>
<td>China</td>
<td>Quantitative: cuasiexperimental</td>
<td>26</td>
<td>Not random</td>
<td>13</td>
<td>13</td>
<td>Not specific</td>
<td>T de Student</td>
<td>Highlights the positive impact of students when using ChatGPT.</td>
</tr>
<tr>
<td>(Chiu T. K., 2024)</td>
<td>China</td>
<td>Quantitative</td>
<td>51</td>
<td>Not random</td>
<td>No</td>
<td>No</td>
<td>Not specific</td>
<td>Topic analysis</td>
<td>Higher education should be prepared for employment in a GenAI-driven society.</td>
</tr>
<tr>
<td>(Bin-Nashwan, Sadallah, &amp; Bouteraa, 2023)</td>
<td>Malaysia</td>
<td>Quantitative</td>
<td>702</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>SEM</td>
<td>Self-esteem and perceived stress in the use of ChatGPT turn out to be positive.</td>
</tr>
<tr>
<td>(Duong, Vu, &amp; Ngo, 2023)</td>
<td>Vietnam</td>
<td>Quantitative</td>
<td>1389</td>
<td>Random stratified</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Multiple lineal regression</td>
<td>Indirectly, they serially increased their actual ChatGPT usage.</td>
</tr>
<tr>
<td>(Gao, Cheah, Lim, &amp; Luo, 2024)</td>
<td>China</td>
<td>Quantitative</td>
<td>376</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>SEM</td>
<td>These findings are of academic importance and have practical implications for educators and students.</td>
</tr>
<tr>
<td>(Essel, Vlachopoulos, Essuman, &amp; Amankwa, 2024)</td>
<td>Ghana</td>
<td>Mix</td>
<td>125</td>
<td>Not random</td>
<td>65</td>
<td>60</td>
<td>Not specific</td>
<td>ANCOVA</td>
<td>ChatGPT influenced critical, reflective, and creative thinking skills.</td>
</tr>
<tr>
<td>(Malik, y otros, 2023)</td>
<td>Indonesia</td>
<td>Quantitative</td>
<td>245</td>
<td>Not random</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>Not specific</td>
<td>The findings indicated a positive reception of AI-powered writing tools.</td>
</tr>
</tbody>
</table>

Table 4. (Continued)
publication, geographic location of studies, type of study, sample size, sample selection criteria, control group, experimental group, validity and reliability of the instruments, as well as the statistical test used and the results obtained.

The findings of this study reveal the geographical distribution of the selected articles according to their country of origin. The significant contribution of articles from China stands out, representing 14.28% of the total sample. There is also a notable presence of contributions from Japan (10.71%), Turkey, Indonesia, Australia, the United Kingdom, and Bulgaria, each representing 7.14%. Other nations, such as Saudi Arabia, the United States, Bangladesh, Morocco, the Philippines, Bolivia, Palestine, Ghana, India, Malaysia, and Vietnam, contributed 3.57% to the analyzed data set. It is crucial to highlight that a geographical bias has been identified in the distribution of documents, evidencing a low representation of publications related to the research topic in Latin America. This disparity is possible because the topic is relatively new or the scarce presence of artificial intelligence laboratories focused on applications related to higher education in regions where such publications have not been recorded.

The results highlight the global importance of research on the application of generative artificial intelligence in higher education. However, it is crucial to recognize that the choice of database could have significantly influenced the observed geographic distribution. Regarding the approach or type of research in the studies reviewed, it is observed that 57.14% of them adopt a quantitative approach, while 17.85% adhere to a qualitative approach. On the other hand, the mixed approach contributes 25.00% of the total. Regarding the sample, 96.42% of the studies have a specified size, while the remaining 3.57% do not. Similarly, 82.14% of the studies show a non-random sample selection, in contrast to 10.71% that present a random selection and 7.14% that do not specify the type of selection. About experimental studies, only 14.28% include both a control group and an experimental group, while 64.28% do not specify the presence of any of these groups.

With the validity and reliability of the research instrument, it is observed that 42.85% of the studies explicitly support the validity and reliability of the instrument used. In comparison, 57.14% do not specify having carried out this evaluation.

After the geographical analysis, it stands out that 64.28% of the articles evaluated come from the Asian continent, while 14.28% originate from Europe. Furthermore, it is observed that 7.14% corresponds to both the oceanic and African continents. In contrast, only 3.57% of the articles come from North America as well as South America. These findings reveal an evident publication bias, reflected in the lack of uniformity in the distribution of publications across the various continents analyzed.

Regarding the methods and statistical tests used, it is observed that 17.85% of the studies used the Structural Equations Model, followed by t-student (10.71%) and ANOVA (7.14%). Other methods include Mann Whitney U Correlation, MAXQDA 2022, Kappa Cohen, Multiple Linear Regression, Partial Least Squares, and Conformal Factor Analysis, each with a 3.57% frequency. In the case of qualitative studies, 10.71% corresponded to the thematic analysis, triangulation, and ethnographic methods, with 3.57% for each. Furthermore, 17.85% of the studies did not specify the statistical test or methodology.

Below, we present evidence to support the impact, time optimization, influence on creativity, and validity of ChatGPT in educational research. Research by Pham et al. (2023) highlights the significant potential of ChatGPT as an effective tool to assist students in higher education. These findings are supported by the studies of Wang and colleagues (2023) as well as the work of Singh, Tayarani-Najaran, and Yaqoob (2023). Furthermore, it is highlighted that ChatGPT’s AI-driven capabilities offer promising opportunities to enhance the learning experience, as confirmed by Bouker (2024).

According to Kayalı et al. (2023), the results of their research indicate that students reported having positive experiences when using ChatGPT in the educational field. This suggests that this tool could play a significant role in improving the learning experience. These findings are further supported by the study of Yilmaz and Yilmaz (2023). Chiu (2024) states that, in general terms, students find motivation
both in the prospect of securing future employment and in the desire to acquire the skills necessary for roles driven by Generative Artificial Intelligence, as confirmed by the study by Kelly, Sullivan, and Strampel (2023). These results offer an overview of three key areas: learning outcomes (Bower et al. 2024), pedagogy (Xu & Correia, 2023), and evaluation (Kiryakova & Angelova, 2023).

The results of the study conducted by Niloy et al. (2024) provide quantitative validation of the qualitative claims and assumptions presented in numerous previous investigations. Specifically, time savings and task management, content inseparability, ease of access, and user-assisted learning have been determined to have a statistically significant and positive impact. These findings align with the results obtained by Chiu (2024), which further reinforces the results.

Indeed, participants stated that ChatGPT provides fast and accurate answers to questions. Furthermore, these responses are highly effective in increasing user satisfaction, as they can quickly and accurately satisfy their needs (Bin-Nashwan, Sadallah, & Bouteraa, 2023). This aspect is reflected in the results of the research carried out by Wang and collaborators (2023), as well as by Malik and his team (2023). Similarly, studies by Talan and Kalinkara (2023) have reported that ChatGPT offers quick answers to questions within seconds. Furthermore, Geerling et al. (2023) have found that ChatGPT provides accurate responses, as expressed by the researchers in their studies.

Habib et al. (2024) highlight the importance of a meticulous approach when integrating AI into creative education. Although AI has the potential to support creative thinking significantly, it has also been observed to impact creative thinking negatively (Cropley, 2023). Therefore, it is essential to reflect on the methods of introducing and applying AI in the educational environment (Kasneci et al., 2023). It has been found that the influence of ChatGPT on critical, reflective, and creative thinking skills coincides with the findings reported by Essel et al. (2024). Furthermore, AI has been found to contribute to the development of divergent thinking, a crucial aspect of the creative process, as evidenced by the results obtained by Habibi et al. (2023).

However, ChatGPT’s ability to process information from text input can reduce the originality of the work, resulting in less creative content (Henriksen, Woo, & Mishra, 2023). The ChatGPT system’s ability to understand human language makes it easy to produce text creatively, such as writing poems, short stories, novels, or other types of writing that can reach the quality equivalent to human work (Shidiq, 2023). This raises concerns about the possible use of ChatGPT in contexts where student creativity is required. Related to this topic, Shorey et al. (2024) highlight the importance of recognizing the legitimate concerns associated with the potential misuse of ChatGPT. As with all technologies we have experienced in the past, ChatGPT is here to stay. Institutions must regulate its use appropriately, adopting artificial intelligence and ChatGPT to optimize their potential while taking necessary precautions when using this technology (Athilingam and He, 2023). Higher education institutions face an urgent challenge to adapt their educational models and teaching methodologies to integrate AI into the teaching-learning process to prepare them for employment in a GenAI-driven society, as Chiu suggests (2024). Along these lines, Habibi et al. (2023) state that higher education institutions (HEIs) could improve the use of ChatGPT by establishing coherent regulations that optimize its application in learning activities.

Through the review of various research, a convergence of results has been observed that suggests a positive impact of ChatGPT on student learning. However, to generalize these findings to a broader population, it is crucial to consider the type of samples used in such research. It has been found that only 10.71% of the studies analyzed used random samples. In comparison, 89.28% opted for non-random samples, such as participants selected by convenience or those who volunteered. This approach can generate biases and errors, both random and systematic, as they do not adequately represent the general population.

Therefore, it is imperative to conduct more experimental research that uses random samples and is more representative of the population. This approach will not only improve the external validity of the results but will also
ensure a more accurate and reliable interpretation of the impact of ChatGPT on student learning. When analyzing the studies included in our review, it is observed that only 3.57% of them correspond to pure experimental studies. As is known, this type of research involves implementing an intervention or treatment that uses a simple random sample, along with a control group and an experimental group. However, since the remaining 96.42% are not purely experimental, a causal relationship cannot be conclusively established or generalized (Campbell and Stanley, 2015).

On the other hand, 10.71% of the studies included in our review present a quasi-experimental design. In this type of research, variables are manipulated, but participants are not randomly assigned to groups. Instead, groups can be formed based on convenience, specific characteristics of the participants, and geographic location, among other criteria. The results are then compared between the groups to determine if the intervention had any effect. Although this type of research is not as rigorous as a pure experimental study, it is considered more robust than a pre-experimental study (Fernandez et al., 2014).

The results reveal that 85.71% of the research did not use a control group, which means that a point of comparison was not available to evaluate the effects of using ChatGPT in higher education. This absence of a control group hinders the validity of the results obtained about the specific impact of the ChatGPT application. In all scientific research, it is crucial to design the study appropriately to avoid errors that could compromise the stated objectives. However, random errors have been identified in most research, which can be attributed to the voluntary selection of participants and the convenience of the researcher in selecting them, which could result in an unrepresentative sample. Furthermore, the validity and reliability of evaluation instruments determine the aspects of solid research. In this sense, only 42.85% of the studies have provided information on the validity and reliability of their instruments.

**Final considerations**

A systematic review was conducted to analyze the impact, time optimization, acceptance, students’ creative process, and research validity when using ChatGPT in higher education. After examining the reviewed articles, it was found that there is a positive impact on the optimization of time, the creative process, and acceptance; however, the validity was not corroborated by a significant percentage. Research has not adequately analyzed the procedures necessary to carry out experimental research, such as the reliability and validity criteria of measurement instruments and the conditions required to carry out experimental research.

This research highlights the need for further study and future analysis to address variations in results based on different prompts or words used with ChatGPT and the potential impact on student satisfaction and effectiveness. Likewise, it opens opportunities for future exploration and improvement in designing and implementing AI-assisted learning systems, ensuring their optimal use and addressing concerns and difficulties that students may face. Therefore, it is recommended that experimental research be carried out with more rigorous criteria in the selection of the sample and in the application of measurement instruments to guarantee the validity and reliability of the results obtained in the research.

**Conflict of interest**

The authors declare that there is no conflict of interest.

**Statement of data consent**

The data generated during the development of this study has been included in the manuscript.

**Contribution statement**

Conceptualization, Methodology, Investigation, Formal Analysis, Writing – review & editing: Daniel Andrade-Girón, William Marín-Rodriguez.

Generative artificial intelligence... Iberoamerican Journal of Science Measurement and Communication

Funding acquisition: Daniel Andrade-Girón, William Marín-Rodriguez, Juana Sandivar-Rosas, Edgardo Carreño-Cisneros, Edgar
Susinarbar-Ramirez, Marcelo Zuñiga-Rojas, Julio Angeles-Morales, Henry Villarreal-Torres.

REFERENCES


Duong, C. D., Vu, T. N., & Ngo, T. V. (2023). Applying a modified technology acceptance model to explain higher education students’


Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: is ChatGPT a blessing or a curse?. *In Frontiers in Education*, 8, 100190. doi:https://doi.org/10.3389/feduc.2023.100190


Revista de práctica de enseñanza y aprendizaje universitario, 20(6), 12. doi:https://doi.org/10.53761/1.20.6.12


Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., & Darwis,


