

SSTSDS-Vol.4. N1. 104

**Flipped classrooms supported by generative artificial intelligence in
higher education**

*Aulas invertidas apoyadas por inteligencia artificial generativa en la
educación superior*

Autores:

Marlene Elizabeth Rengifo Obando

Universidad Indoamérica

Quito – Ecuador

marlene.rengifo@uti.edu.ec

<https://orcid.org/0009-0009-2007-6444>

Juan Gabriel Vera Alcívar

Unidad Educativa Soldado Monge

El Empalme – Ecuador

soldadomongeingles@gmail.com

<https://orcid.org/0009-0009-8174-5680>

Carolina de Lourdes Arcentales Hidalgo

Investigadora independiente

Quito – Ecuador

carolina.arcentales@docentes.educacion.edu.ec

<https://orcid.org/0009-0004-7248-411X>

Karen Sabrina Brito Ordóñez

Investigadora independiente

Riobamba – Ecuador

sabrii.bo1580@gmail.com

<https://orcid.org/0009-0004-0690-848X>

Poleth Alejandra Mendieta Sinche

Investigadora independiente

Cuenca – Ecuador

poleth.mendieta@gmail.com

<https://orcid.org/0000-0002-0848-1270>

Autor de correspondencia: Marlene Elizabeth Rengifo Obando, soldadomongeingles@gmail.com

Enviado: 13-05-2026

Aceptado: 03-08-2026

Revisado: 28-06-2026

Publicado: 05-08-2026



Cómo citar este artículo:

Rengifo Obando, M. E., Vera Alcívar, J. G., CaroArcentales Hidalgo, C. de L., Brito Ordóñez, K. S., & Mendieta Sinche, P. A. (2026). Aulas invertidas apoyadas por inteligencia artificial generativa en la educación superior. *Sage Sphere of Technology, Sciences, Discoveries And Society*, 4(1), 1-28. <https://doi.org/10.63688/d9tvzf37>

© 2026; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (<https://creativecommons.org/licenses/by/4.0>) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea correctamente citada.



ABSTRACT

The generative artificial intelligence-supported flipped classroom combines autonomous pre-class preparation with active, collaborative, and teacher-guided classroom activities through tools such as ChatGPT, Gemini, and Copilot. This study aimed to analyze the contribution of this pedagogical approach to learning outcomes in higher education. An integrative literature review was conducted through a systematic search process based on the identification, screening, eligibility, and inclusion stages established by PRISMA 2020. The literature search covered the period from 2020 to 2026, and 15 publications that met the eligibility criteria were included. The findings were organized into four thematic categories: academic achievement and learning performance; student engagement and active participation; self-regulated learning and higher-order thinking; and pedagogical, technological, and ethical implementation conditions. The results indicate that generative artificial intelligence can strengthen the flipped classroom by providing personalized explanations, immediate feedback, pre-class preparation, active participation, and learning support. However, its effectiveness depends on instructional design, student preparation, technological access, and teacher guidance. Challenges related to academic integrity, privacy, dependence on AI, and unequal access were also identified. It is concluded that AI-supported flipped classrooms can contribute to higher education when implemented as a structured pedagogical model that promotes the critical verification of AI-generated information and active student participation under continuous teacher guidance.

Keywords: generative artificial intelligence; flipped classroom; higher education; student engagement; self-regulated learning.

RESUMEN

El aula invertida apoyada por inteligencia artificial generativa combina la preparación autónoma previa con actividades presenciales activas, colaborativas y orientadas por el docente mediante herramientas como ChatGPT, Gemini y Copilot. Este estudio tuvo como objetivo analizar la contribución de este enfoque pedagógico a los resultados de aprendizaje en la educación superior. Se realizó una revisión integrativa de la literatura mediante un proceso de búsqueda sistemática basado en las etapas de identificación, cribado, elegibilidad e inclusión establecidas por PRISMA 2020. La búsqueda bibliográfica abarcó el periodo 2020–2026 y se incluyeron 15 publicaciones que cumplieron los criterios de elegibilidad. Los hallazgos se organizaron en cuatro categorías temáticas: rendimiento académico y desempeño del aprendizaje; participación activa del estudiante; aprendizaje autorregulado y pensamiento de orden superior; y condiciones pedagógicas, tecnológicas y éticas de implementación. Los resultados muestran que la inteligencia artificial generativa puede fortalecer el aula invertida mediante explicaciones personalizadas, retroalimentación inmediata, preparación previa, participación activa y apoyo al aprendizaje. No obstante, su efectividad depende del diseño instruccional, la preparación del estudiante, el acceso tecnológico y la orientación docente. Asimismo, se identificaron desafíos relacionados con la integridad académica, la privacidad, la dependencia de la IA y el acceso desigual. Se concluye que el aula invertida apoyada por IA puede contribuir a la educación superior cuando se implementa como un modelo pedagógico estructurado que promueva la verificación crítica de la información y la participación activa bajo la mediación continua del docente.

Palabras clave: inteligencia artificial generativa; aula invertida; educación superior; participación estudiantil; aprendizaje autorregulado.



1. INTRODUCTION

Higher education requires students to develop critical thinking, problem-solving, collaboration, creativity, digital skills, and independent learning. However, traditional teaching often uses most classroom time for explanations, leaving fewer opportunities for participation, practical activities, and feedback. In response to this limitation, flipped learning has gained relevance as a student-centered methodology that reorganizes the traditional sequence of instruction by transferring initial content acquisition outside the classroom and using class time for discussion, collaboration, problem-solving, application of knowledge, and higher-order learning activities.

Research has shown positive results from this methodology. Strelan et al. (2020) found that flipped instruction produced a moderate improvement in student performance because it created more opportunities for structured active learning and problem-solving. Similarly, Baig and Yadegaridehkordi (2023) concluded that flipped classrooms can improve engagement, autonomy, collaboration, and academic achievement. Bredow et al. (2021) also analyzed evidence from higher education and reported positive effects of flipped learning on academic performance, interpersonal development, and student satisfaction. Nevertheless, these studies show that the effectiveness of the flipped classroom depends not only on the use of digital resources, but also on instructional design, student preparation, technological access, teacher support, and the connection between pre-class and in-class activities.

The development of generative artificial intelligence has created new possibilities for flipped learning. Tools such as ChatGPT, Gemini, and Copilot can generate explanations, summaries, examples, questions, learning materials, and immediate feedback. Before class, students may use these tools to clarify difficult concepts, summarize readings, create practice questions, and receive personalized explanations. During class, they can compare AI-generated responses, identify errors, discuss alternative solutions, solve problems, and complete collaborative activities. After class, AI can provide additional practice, feedback, and reflection activities. In this sense, generative AI may complement flipped learning by supporting students during the pre-class stage and strengthening classroom activities through personalized guidance, formative feedback, and adaptive learning support.



Several studies have examined the educational potential of generative AI. Kasneci et al. (2023) indicate that generative AI can support personalized learning, interaction, feedback, and educational content creation. Lo (2023) found that ChatGPT can support education by generating explanations, learning materials, feedback, and personalized assistance. Bond et al. (2024) also reported that artificial intelligence in higher education can support personalized learning, student engagement, assessment, and academic decision-making. However, these authors emphasized that AI implementation requires ethical principles, collaboration among educators and researchers, and rigorous evaluation of its educational effects.

In the context of flipped learning, Apaza Calsin et al. (2023) described the use of ChatGPT within a flipped classroom for teaching basic programming concepts. Their study showed that students need to formulate appropriate questions and critically examine the answers generated by AI. Essel et al. (2024) also found that structured ChatGPT activities supported critical, creative, and reflective thinking among undergraduate students. These findings suggest that AI can contribute to higher-order thinking when students are required to evaluate information, explain their reasoning, and develop their own conclusions instead of simply copying generated answers.

In programming education, Huesca et al. (2024) found that students who participated in a ChatGPT-supported flipped classroom obtained better learning results than those who followed a flipped model based mainly on instructional videos. The authors associated these improvements with the immediate and personalized explanations provided by the AI tool. Likewise, Lee et al. (2024) demonstrated that generative AI was more effective when it provided guidance, questions, and hints rather than complete answers. Their findings showed improvements in self-regulated learning, higher-order thinking, and knowledge construction, suggesting that AI should function as a learning assistant that encourages reflection and progressive problem-solving.

Generative AI may also contribute to student engagement, motivation, and digital competence. Yavuz et al. (2025) found that AI-supported flipped classroom activities increased students' motivation, participation, and artificial intelligence literacy. However, participants also reported difficulties related to workload, time management, and the need to critically evaluate AI-generated information. More recent research has reported



similar findings in different educational areas. Cabrera-Solano et al. (2026) found improvements in linguistic accuracy, writing, feedback, self-regulation, and understanding of pedagogical concepts among pre-service EFL teachers who used ChatGPT, Gemini, and Copilot within a flipped classroom. H. Liu et al. (2026) reported improvements in academic achievement and emotional well-being in a university mental health course. Additionally, Y. J. Liu et al. (2026) found that an AI-enhanced flipped classroom improved preparation, assignment accuracy, and classroom participation in pharmacology education, although it did not produce superior results in every learning area.

Despite these benefits, AI-supported flipped learning also presents pedagogical and ethical challenges. Students may receive inaccurate, incomplete, or biased information, become dependent on AI, submit work without understanding it, or use generated content dishonestly. Other concerns include privacy, unequal technological access, teacher preparation, academic integrity, assessment reliability, and increased workload. Farrokhnia et al. (2024) identified personalized responses, immediate interaction, access to information, and reduced teacher workload as important strengths of ChatGPT, but also emphasized weaknesses and threats associated with inaccurate answers, excessive dependence, academic dishonesty, and limited critical evaluation. Therefore, students should verify AI responses, consult academic sources, explain their reasoning, and clearly report how they used AI.

For this reason, the teacher's role remains central in AI-supported flipped classroom models. Rather than being replaced by technology, instructors act as pedagogical mediators who design meaningful learning tasks, guide ethical and responsible use of AI, validate the quality of generated content, and help students connect digital support with disciplinary knowledge. Teacher mediation is especially important because generative AI can provide immediate assistance, but it cannot guarantee that students understand, evaluate, or apply information critically. Therefore, the educational value of AI depends not only on the tool itself, but also on how it is integrated into the learning design, how students are guided to use it, and how its outputs are discussed, questioned, and validated within the classroom.

Overall, research published between 2020 and 2026 suggests that generative AI can strengthen flipped classroom practices by improving preparation, feedback, participation, self-regulated learning, higher-order thinking, artificial intelligence literacy, emotional well-being, and academic performance. Nevertheless, AI should be used as a learning



support tool and not as a replacement for teachers or independent thinking. Although flipped learning and generative AI have been widely discussed as separate areas of educational innovation, their combined use in higher education still requires deeper analysis because the available evidence remains methodologically diverse and includes empirical studies, systematic reviews, meta-analyses, rapid reviews, and conceptual publications.

Accordingly, this study aims to analyze how generative artificial intelligence-supported flipped classrooms influence academic achievement, student engagement, self-regulated learning, and higher-order thinking in higher education. The guiding research question is: How does the integration of generative artificial intelligence into the flipped classroom influence learning outcomes, student engagement, and self-regulated learning in higher education?

2. METHODOLOGY

This study was conducted as an integrative literature review with a systematic search process to analyze the effects of generative artificial intelligence-supported flipped classrooms in higher education. This design was selected because it allows empirical studies, experimental and quasi-experimental research, systematic reviews, meta-analyses, and conceptual publications to be examined within a common analytical framework. Integrative reviews facilitate the synthesis of different forms of evidence and contribute to a broader understanding of an emerging educational phenomenon (Lubbe et al., 2020).

A qualitative descriptive-analytical approach was used to organize and interpret the selected publications. The analysis considered academic achievement, student engagement, active participation, self-regulated learning, higher-order thinking, personalized feedback, artificial intelligence literacy, emotional well-being, and the pedagogical and ethical conditions associated with AI-supported flipped learning.

The literature search covered publications from 2020 to 2026. The final corpus consisted of 15 publications related to flipped classroom methodology, generative artificial intelligence, or the integration of both approaches in higher education. The selected evidence included empirical studies, experimental and quasi-experimental research, systematic reviews, meta-analyses, rapid reviews, and conceptual analyses.

Information Sources and Search Strategy



The bibliographic search was conducted in Scopus, Web of Science, ScienceDirect, SpringerLink, Taylor & Francis Online, SAGE Journals, ERIC, and Google Scholar.

Searches were performed in title, abstract, and keyword fields whenever these options were available. The search strategy combined terms related to generative artificial intelligence, flipped learning, and higher education through the Boolean operators AND and OR.

The general search equation was:

(“generative artificial intelligence” OR “generative AI” OR “ChatGPT” OR “large language models”) AND (“flipped classroom” OR “flipped learning” OR “inverted classroom”) AND (“higher education” OR “university students” OR “college students”)

Additional searches incorporated learning-related outcomes:

(“generative AI” OR “ChatGPT”) AND (“flipped classroom” OR “flipped learning”) AND (“academic achievement” OR “student engagement” OR “self-regulated learning” OR “higher-order thinking” OR “AI literacy”)

The equations were adapted to the fields, operators, and filters available in each database. The references were organized in a Microsoft Excel matrix. Duplicate records were identified by comparing authorship, title, publication year, journal, and DOI.

Although the search strategy prioritized studies combining generative artificial intelligence, flipped classroom, and higher education, high-quality systematic reviews, meta-analyses, and conceptual studies addressing either flipped learning or generative artificial intelligence were also included when they provided essential theoretical foundations or contextual evidence for interpreting the findings. Consequently, the thematic synthesis was based primarily on studies that directly examined AI-supported flipped classrooms, whereas broader publications on flipped learning or generative AI were used exclusively to contextualize the results and strengthen the interpretation of the evidence.

Eligibility Criteria

Two complementary categories of publications were included in this review. The first comprised empirical studies that directly examined the integration of generative artificial intelligence into flipped classrooms in higher education and constituted the primary evidence for the thematic synthesis. The second included high-quality systematic reviews, meta-analyses, and conceptual studies on flipped learning or generative artificial intelligence that provided theoretical and contextual support for interpreting the findings. These



broader publications were incorporated to strengthen the conceptual framework and discussion but were not considered direct evidence of AI-supported flipped classroom implementation.

Evidence Classification and Synthesis

To minimize the risk of duplicating evidence, the included publications were classified into three complementary categories: (1) primary empirical studies (experimental, quasi-experimental, and case studies) that directly evaluated AI-supported flipped classroom interventions; (2) secondary evidence (systematic reviews, meta-analyses, and rapid reviews) that synthesized previous research; and (3) conceptual publications that provided theoretical or methodological perspectives. During the thematic synthesis, findings from each category were interpreted according to their methodological purpose. Primary studies constituted the main source of evidence regarding educational outcomes, whereas secondary and conceptual publications were used to contextualize, compare, and interpret the empirical findings rather than to increase the weight of the evidence. This approach reduced the possibility of counting the same empirical evidence more than once.

Publication Selection Process

The selection process followed the identification, screening, eligibility, and inclusion phases established by the PRISMA 2020 statement, which provides guidance for transparently reporting how evidence is identified, evaluated, excluded, and included in systematic reviews (Page et al., 2021).

The database search identified 142 records. After removing 37 duplicates and five records for other reasons, 100 publications remained for title and abstract screening. Sixty-two records were excluded because they did not sufficiently address the research topic. Of the 38 reports sought for retrieval, three could not be accessed, resulting in 35 full-text publications assessed for eligibility. Twenty publications were excluded because they were unrelated to higher education, did not examine flipped learning or generative AI, focused exclusively on technical AI development, or presented insufficient methodological information. Consequently, 15 publications were included in the final review.

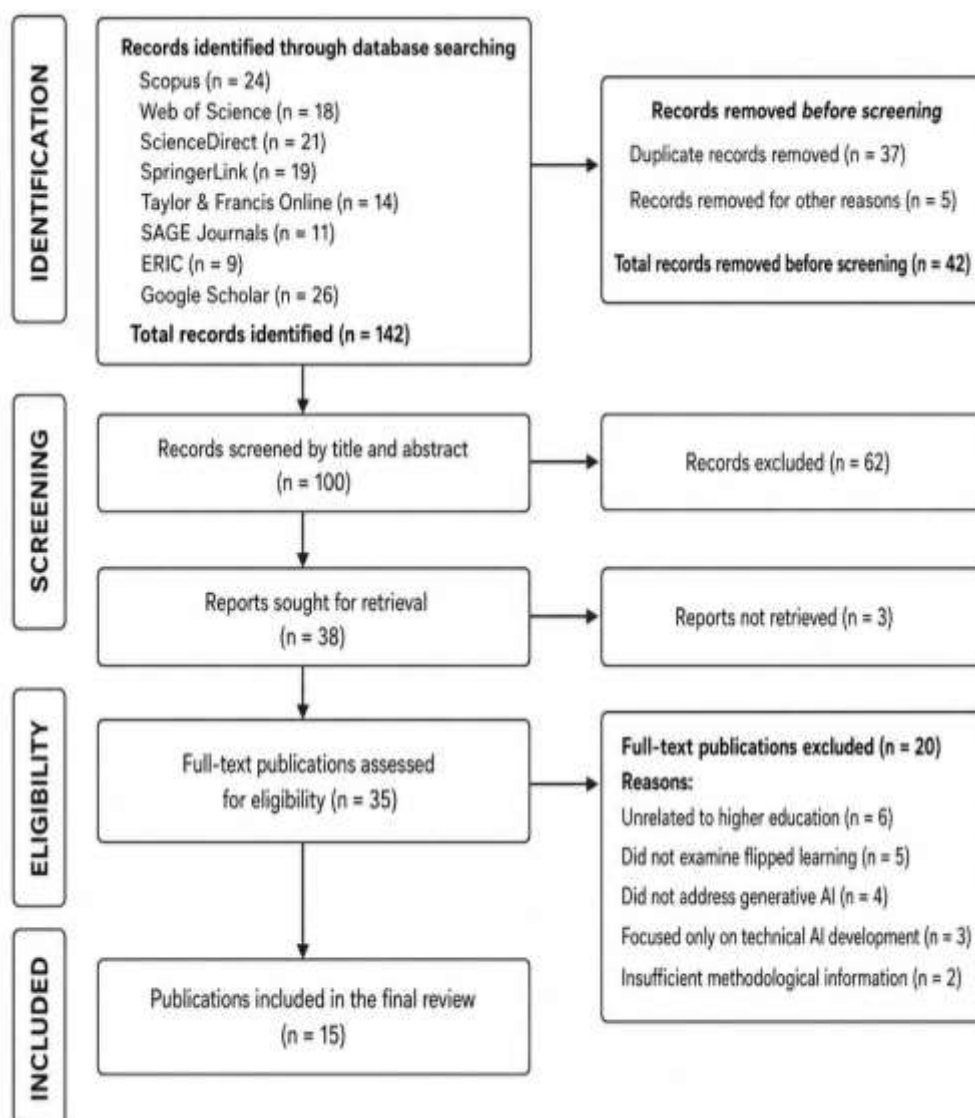
First, the records identified through the selected databases were compiled. Duplicate publications were removed before screening. The remaining titles and abstracts were evaluated according to the research topic and eligibility criteria. Potentially



relevant publications were subsequently reviewed in full text. Publications that did not meet the inclusion criteria were excluded, and 15 studies were retained in the final corpus.

Figure 1

PRISMA 2020 Flow Diagram of the Publication-Selection Process



Note. The publication-selection process followed the identification, screening, eligibility, and inclusion stages established by the PRISMA 2020 statement. A total of 142 records were identified through eight academic information sources. After removing 42 duplicate or otherwise ineligible records, 100 publications were screened by title and abstract. Of these, 62 records were excluded, leaving 38 reports sought for retrieval. Three reports could not be retrieved; therefore, 35 full-text publications were assessed for eligibility. Of these, 20 were



excluded for not meeting the established criteria. Finally, 15 publications published between 2020 and 2026 were included in the integrative review.

A standardized matrix was used to extract the following information:

- Author and publication year.
- Type of publication or research design.
- Educational context.
- Participants or analyzed corpus.
- Flipped classroom strategy.
- Generative AI tool or technological resource.
- Learning outcomes examined.
- Main findings.
- Reported limitations.
- DOI or bibliographic identifier.
- Overall contribution to AI-supported flipped learning.

The evidence was analyzed through thematic categorization. Thematic analysis allows recurrent patterns to be identified, organized, interpreted, and presented as meaningful categories (Braun & Clarke, 2021).

The findings were organized into four thematic categories:

1. Academic achievement and learning performance.
2. Student engagement and active participation.
3. Self-regulated learning and higher-order thinking.
4. Pedagogical, technological, and ethical implementation conditions.

The categories were not mutually exclusive because some publications addressed more than one academic or educational variable. A new statistical meta-analysis was not conducted because the selected publications differed in design, participants, academic disciplines, intervention duration, AI tools, learning outcomes, and assessment procedures. Therefore, the evidence was synthesized descriptively and thematically.

Methodological Quality Assessment

A common eight-criterion framework was used to evaluate methodological and reporting elements across the different publication types:

C1: Clarity of objectives.



C2: Adequacy of the theoretical foundation.

C3: Methodological transparency.

C4: Appropriateness of the research design.

C5: Description of participants, corpus, or evidence.

C6: Clarity of the analysis.

C7: Correspondence between findings and conclusions.

C8: Acknowledgment of limitations.

The methodological quality assessment was conducted using a differentiated approach according to the type of publication included in the review. Given the methodological diversity of the corpus, the appraisal considered the specific design and purpose of each publication. Systematic reviews and meta-analyses, empirical studies, mixed-methods research, and conceptual publications were interpreted according to their methodological nature rather than being judged through a single uniform scale.

The author-developed eight-criterion matrix was retained only as a descriptive synthesis tool to organize and compare reporting quality across the included publications. Therefore, the scores derived from this matrix should not be interpreted as a formal risk-of-bias assessment or as a fully validated quality appraisal instrument.

Each criterion was scored as follows:

0: The criterion was not met.

1: The criterion was partially met.

2: The criterion was fully met.

The maximum possible score was 16 points. The quality percentage was calculated by dividing the score obtained by 16 and multiplying it by 100.

Publications obtaining 80% or more were classified as high quality, those obtaining between 60% and 79% as moderate quality, and those obtaining less than 60% as low quality.

Table 1

Descriptive Reporting-Quality Assessment of the Included Publications

Study	Type of publication	Criteria met	Score	Descriptive reporting quality
Strelnan et al. (2020)	Meta-analysis	8/8	16/16	High reporting quality



Study	Type of publication	Criteria met	Score	Descriptive reporting quality
Bredow et al. (2021)	Meta-analysis	8/8	16/16	High reporting quality
Kasneci et al. (2023)	Conceptual / analytical article	7/8	14/16	High reporting quality
Lo (2023)	Rapid review / review article	7/8	14/16	High reporting quality
Baig and Yadegaridehkordi (2023)	Systematic review	8/8	16/16	High reporting quality
Apaza Calsin et al. (2023)	Empirical study	7/8	14/16	High reporting quality
Farrokhnia et al. (2024)	Systematic review	8/8	16/16	High reporting quality
Bond et al. (2024)	Review / conceptual article	7/8	14/16	High reporting quality
Essel et al. (2024)	Empirical study	7/8	14/16	High reporting quality
Huesca et al. (2024)	Empirical / experimental study	8/8	16/16	High reporting quality
Lee et al. (2024)	Empirical study	7/8	14/16	High reporting quality
Yavuz et al. (2025)	Empirical study	7/8	14/16	High reporting quality
Cabrera-Solano et al. (2026)	Empirical study	8/8	16/16	High reporting quality
H. Liu et al. (2026)	Empirical study	7/8	14/16	High reporting quality
<i>Y. J. Liu et al. (2026)</i>	<i>Empirical study</i>	<i>7/8</i>	<i>14/16</i>	<i>High reporting quality</i>

Note. C1 = clarity of objectives; C2 = adequacy of the theoretical foundation; C3 = methodological transparency; C4 = appropriateness of the research design; C5 = description of participants, corpus, or evidence; C6 = clarity of the analysis; C7 = correspondence between findings and conclusions; and C8 = acknowledgment of limitations. The column “Criteria met” summarizes the number of criteria rated as fully achieved. The percentages represent descriptive reporting quality and should not be interpreted as definitive evidence of low risk of bias or as proof that all studies have the same level of methodological rigor. The high scores reflect the degree to which each publication reported objectives, methods, evidence, analysis, conclusions, and limitations, but the interpretation of quality must consider the specific design of each study.

The descriptive assessment showed that the 15 included publications obtained reporting-quality scores ranging from 87.5% to 100%. Six publications obtained the



maximum score of 16 out of 16 points: Strelan et al. (2020), Bredow et al. (2021), Baig and Yadegaridehkordi (2023), Farrokhnia et al. (2024), Huesca et al. (2024), and Cabrera-Solano et al. (2026). These publications fully met the eight descriptive reporting criteria considered in the assessment.

The remaining nine publications obtained 14 out of 16 points, equivalent to 87.5%: Kasneci et al. (2023), Lo (2023), Apaza Calsin et al. (2023), Bond et al. (2024), Essel et al. (2024), Lee et al. (2024), Yavuz et al. (2025), H. Liu et al. (2026), and Y. J. Liu et al. (2026). Although these publications also fell within the high reporting-quality category, one or more criteria were only partially met or were not reported with the same level of detail.

These results should be interpreted cautiously. The scores reflect compliance with the author-developed descriptive reporting criteria and should not be interpreted as a formal risk-of-bias assessment or as evidence that all publications had equivalent methodological rigor. Differences in research design, participants, analytical procedures, intervention duration, and publication type must also be considered when interpreting the results.

Review Process and Protocol

Publication identification, screening, full-text assessment, data extraction, thematic categorization, and methodological quality evaluation were conducted by one reviewer. The extracted information was reviewed using the full texts of the selected publications. When a methodological characteristic was not explicitly reported by the original authors, it was recorded as “Not explicitly reported” without making assumptions or presenting alternative interpretations.

Because no second independent reviewer participated, no inter-rater agreement coefficient was calculated. These methodological conditions were considered during the interpretation of the findings. Since study selection, data extraction, and quality appraisal were conducted by a single reviewer, the results may be affected by selection bias, subjective interpretation, or unintentional omissions. Therefore, the thematic findings and quality classifications should be understood as provisional and descriptive rather than as definitive conclusions supported by independent reviewer agreement. The review was not prospectively registered, and no public protocol was published before the search. Nevertheless, the databases, search equations, eligibility criteria, selection stages, extraction variables, analytical categories, and quality criteria were documented to improve transparency and reproducibility.



3. RESULTS

The thematic synthesis identified four main categories related to the integration of generative artificial intelligence into flipped classroom experiences in higher education: academic achievement and learning performance; student engagement and active participation; self-regulated learning and higher-order thinking; and pedagogical, technological, and ethical implementation conditions. These categories were not supported equally across the included publications. Some were based on direct empirical evidence, while others were mainly supported by secondary or conceptual publications.

Academic achievement was reported in several empirical studies, particularly in programming, English-language teacher education, mental health education, and pharmacology. Positive effects were associated with immediate explanations, structured pre-class preparation, personalized feedback, and opportunities for active application during class. However, the evidence was not uniform, since some studies reported stronger effects in preparation and participation than in all dimensions of academic performance.

Student engagement was another recurrent category. Several publications indicated that AI-supported flipped learning increased participation, motivation, interaction, and classroom involvement. These improvements were mainly observed when students used AI tools to prepare before class, generate ideas, ask questions, and participate in collaborative activities. Nevertheless, engagement was also influenced by workload, time management, digital skills, and the clarity of teacher guidance.

Self-regulated learning appeared as a relevant outcome in studies where students used generative AI to plan, review, practice, and monitor their learning. AI tools supported autonomous preparation and individualized assistance, but this effect depended on students' ability to formulate appropriate prompts, evaluate responses, and use feedback critically.

Higher-order thinking was reported when AI was used not only to provide answers but also to stimulate questioning, comparison, reflection, error detection, and problem-solving. Studies that required students to analyze or critique AI-generated content showed stronger connections with critical and creative thinking than those focused only on receiving explanations.



Finally, the review identified several pedagogical and ethical risks. These included overdependence on AI-generated responses, inaccurate information, superficial learning, academic dishonesty, privacy concerns, unequal access to technology, and the need for teacher training. These risks were especially emphasized in secondary and conceptual publications, although empirical studies also reported practical challenges related to workload, student preparation, and the need for critical evaluation.

Table 2

Primary Empirical Studies on AI-Supported Flipped Learning

Study	Context / discipline	AI tool or support	Flipped-learning use	Main outcome
Apaza Calsin et al. (2023)	Basic programming	ChatGPT	Support for pre-class preparation and classroom problem-solving	Helped students formulate questions and critically examine AI-generated answers
Essel et al. (2024)	Undergraduate education	ChatGPT	Structured AI-supported activities within flipped learning	Supported critical, creative, and reflective thinking
Huesca et al. (2024)	Programming education	ChatGPT	AI-supported flipped classroom compared with video-based flipped learning	Improved learning results through immediate and personalized explanations
Lee et al. (2024)	Higher education	Generative AI guidance	AI provided hints, questions, and guidance instead of complete answers	Improved self-regulated learning, higher-order thinking, and knowledge construction
Yavuz et al. (2025)	Higher education	Generative AI tools	AI-supported flipped classroom activities	Increased motivation, participation, and AI literacy, but showed workload and time-management challenges
Cabrera-Solano et al. (2026)	English-language teacher education	ChatGPT, Gemini, and Copilot	AI-supported flipped classroom for language and pedagogical learning	Improved linguistic accuracy, writing, feedback, self-regulation, and pedagogical understanding



Study	Context / discipline	AI tool or support	Flipped-learning use	Main outcome
H. Liu et al. (2026)	University mental health education	Generative AI support	AI-enhanced flipped classroom	Improved academic achievement and emotional well-being
Y. J. Liu et al. (2026)	Pharmacology education	Generative AI support	AI-enhanced flipped classroom for preparation and classroom participation	Improved preparation, assignment accuracy, and participation, although effects were not superior in all areas

Note. This table includes primary empirical studies that directly examined the implementation of generative artificial intelligence within flipped classroom experiences. The table was reduced to essential comparison categories to improve readability.

Table 3

Secondary and Conceptual Publications Used for Contextual Support

Study	Type of publication	Main focus	Contribution to the review
Strelan et al. (2020)	Meta-analysis	Effects of flipped instruction on student performance	Provides evidence that flipped learning can improve academic performance through active learning and problem-solving
Bredow et al. (2021)	Meta-analysis	Flipped learning in higher education	Supports the discussion of performance, interpersonal development, and student satisfaction
Kasneci et al. (2023)	Conceptual / analytical article	Opportunities and challenges of generative AI in education	Supports the discussion of personalized learning, feedback, content creation, and ethical challenges
Lo (2023)	Rapid review / review article	ChatGPT in education	Provides contextual evidence on explanations, learning materials, feedback, and personalized assistance
Baig and Yadegaridehkordi (2023)	Systematic review	Flipped classroom benefits and limitations	Supports the analysis of engagement, autonomy, collaboration, and academic achievement



Farrokhnia et al. (2024)	Systematic review	Strengths, weaknesses, opportunities, and threats of ChatGPT	Supports the discussion of AI benefits, dependence, misinformation, academic dishonesty, and critical evaluation
Bond et al. (2024)	Review / conceptual article	Artificial intelligence in higher education	Provides contextual support on personalized learning, engagement, assessment, and ethical implementation

Note. These publications were not treated as direct evidence of AI-supported flipped classroom implementation. They were used to contextualize the findings, compare trends, and support the interpretation of pedagogical benefits, risks, and institutional implications. The thematic synthesis presented below is organized according to the methodological nature of the included evidence. Primary empirical studies that directly examined the integration of generative artificial intelligence into flipped classrooms in higher education provided the principal evidence for the analysis. Systematic reviews, meta-analyses, rapid reviews, and conceptual publications were incorporated as contextual evidence to support the interpretation and discussion of the findings but were not considered direct evidence of AI-supported flipped classroom implementation. This classification was adopted to minimize the risk of duplicating evidence derived from the same primary studies.

Academic Achievement and Learning Performance

The reviewed evidence showed a generally favorable relationship between flipped learning and academic achievement. The two meta-analyses indicated that flipped instruction produced positive effects on student performance because it increased opportunities for active learning, problem-solving, and classroom application (Strelan et al., 2020; Bredow et al., 2021).

Studies that directly incorporated generative AI also reported positive outcomes. Huesca et al. (2024) found that students using a ChatGPT-supported flipped strategy achieved higher learning gains than those using a conventional video-based model. Similarly, Apaza Calsin et al. (2023) showed that ChatGPT could support the explanation and application of programming concepts, although students still needed to test and justify the generated solutions.



The findings suggest that AI does not improve achievement automatically. Its value depends on its connection with learning objectives, pre-class preparation, classroom activities, teacher supervision, and assessment procedures.

Student Engagement and Active Participation

The selected publications indicated that AI-supported flipped classrooms may promote participation by providing students with explanations, questions, examples, and feedback before class. This preparation may allow classroom time to be used for discussion, collaborative problem-solving, projects, and practical application.

Yavuz et al. (2025) reported positive student perceptions, greater interest, and opportunities to develop artificial intelligence literacy. Baig and Yadegaridehkordi (2023) similarly concluded that flipped learning can strengthen engagement, autonomy, and collaboration when pre-class materials are clearly connected with classroom activities.

Nevertheless, participation may decrease when students fail to complete preparatory activities, encounter technological difficulties, or perceive the workload as excessive. Therefore, engagement depends on accessible resources, clear instructions, preparation monitoring, and purposeful classroom tasks.

Self-Regulated Learning and Higher-Order Thinking

AI-supported flipped learning requires students to organize their time, review materials independently, formulate questions, evaluate information, and prepare for active classroom participation. Lee et al. (2024) found that a guidance-based ChatGPT assistant improved self-regulated learning, higher-order thinking, and knowledge construction. The guided system was more effective because students were required to propose an initial response before receiving hints and feedback.

Essel et al. (2024) also reported that structured ChatGPT-supported activities positively influenced critical, creative, and reflective thinking. These findings demonstrate that higher-order thinking is more likely to develop when AI encourages analysis and reflection instead of immediately providing final answers.

Overall, generative AI can function as a cognitive scaffold. However, students must remain responsible for verifying information, explaining their reasoning, and constructing their own conclusions.

Personalized Feedback, AI Literacy, and Emotional Well-Being



Generative AI provides immediate and adaptable responses that may complement teacher feedback. Cabrera-Solano et al. (2026) found that the integration of ChatGPT, Gemini, and Copilot supported writing, linguistic accuracy, feedback, and self-regulation among pre-service EFL teachers.

Yavuz et al. (2025) reported that students developed more positive perceptions and greater AI literacy through the use of different artificial intelligence tools. H. Liu et al. (2026) also identified improvements in academic achievement and emotional well-being, suggesting that flexible and personalized support may influence both cognitive and affective dimensions of learning.

Nevertheless, personalized feedback must be verified because generative AI can produce incomplete, inaccurate, or misleading responses. Teacher supervision remains necessary to ensure that feedback supports rather than replaces learning.

Pedagogical, Technological, and Ethical Implementation Conditions

The reviewed literature emphasized that the effectiveness of AI-supported flipped learning depends on pedagogical design rather than on the simple availability of technology. Kasneci et al. (2023) identified personalization, feedback, and content generation as important opportunities, but also warned about bias, inaccurate information, dependence, privacy, and academic integrity.

Lo (2023) similarly concluded that ChatGPT could support teaching and learning while creating concerns regarding assessment and responsible use. Farrokhnia et al. (2024) identified immediate interaction and personalized responses as strengths, but emphasized threats related to misinformation, excessive dependence, and academic dishonesty.

Bond et al. (2024) argued that AI adoption in higher education requires ethical principles, collaboration, methodological rigor, and evidence-based decision-making. The findings therefore indicate that universities should develop clear policies regarding verification, attribution, privacy, acceptable assistance, and academic integrity.

Overall, the reviewed evidence suggests that generative artificial intelligence can strengthen flipped classrooms by supporting preparation, participation, personalized feedback, self-regulation, higher-order thinking, and academic performance. However, its effectiveness depends on instructional design, teacher guidance, technological access, ethical use, and students' ability to critically evaluate AI-generated information.



4. DISCUSSION

The findings of this integrative review suggest that generative artificial intelligence can strengthen flipped classroom experiences in higher education, although its effects are neither automatic nor homogeneous. The empirical evidence indicates that positive outcomes depend on how AI is incorporated into the learning sequence, the type of assistance provided, the academic discipline, and the degree of teacher mediation. Studies conducted in programming, language education, mental health education, and pharmacology reported improvements in preparation, participation, feedback, academic performance, or self-regulated learning; however, the magnitude and nature of these effects differed across contexts (Apaza Calsin et al., 2023; Huesca et al., 2024; Cabrera-Solano et al., 2026; H. Liu et al., 2026; Y. J. Liu et al., 2026).

Comparison among the primary empirical studies shows that the role assigned to generative AI was a central factor. Huesca et al. (2024) reported better learning results when ChatGPT was incorporated into a flipped classroom than when students mainly used instructional videos. Similarly, Cabrera-Solano et al. (2026) identified improvements in linguistic accuracy, writing, feedback, self-regulation, and understanding of pedagogical concepts. In contrast, Y. J. Liu et al. (2026) found benefits in preparation, assignment accuracy, and classroom participation, but the AI-enhanced model did not produce superior results in every learning area. These differences suggest that generative AI may improve particular dimensions of learning without necessarily producing uniform gains across all academic outcomes.

The evidence also indicates that academic performance should be differentiated from participation and preparation. Improvements in classroom involvement or completion of preparatory activities do not automatically demonstrate deeper disciplinary learning. Yavuz et al. (2025) reported increased motivation, participation, and AI literacy, but also identified workload and time-management challenges. Likewise, Y. J. Liu et al. (2026) observed stronger preparation and participation without superior outcomes in all evaluated areas. Therefore, the positive effects of AI-supported flipped learning should be interpreted according to the specific variable measured rather than treated as a single general improvement in learning.



Self-regulated learning and higher-order thinking also require separate interpretation. Lee et al. (2024) found that generative AI was more effective when it provided hints, questions, and guidance instead of complete answers. This guided approach improved self-regulated learning, higher-order thinking, and knowledge construction because students were required to formulate an initial response and progressively revise their reasoning. Essel et al. (2024) similarly reported that structured ChatGPT-supported activities promoted critical, creative, and reflective thinking. Apaza Calsin et al. (2023) emphasized that students needed to formulate appropriate questions and critically evaluate AI-generated answers. Together, these studies suggest that higher-order thinking is more likely to develop when students analyze, compare, justify, or correct AI-generated information rather than simply receive completed responses.

The comparison between primary, secondary, and conceptual evidence is also relevant. The primary empirical studies provide direct information about the effects of AI-supported flipped classrooms on students and learning outcomes. By contrast, the meta-analyses and systematic reviews offer broader evidence about flipped learning or generative AI separately. Strelan et al. (2020) and Bredow et al. (2021) found that flipped learning generally improves academic performance by increasing opportunities for active learning and problem-solving. Baig and Yadegaridehkordi (2023) also identified positive effects on engagement, autonomy, collaboration, and achievement. However, these publications do not directly demonstrate the effectiveness of generative AI within flipped classroom interventions and should therefore be used as contextual rather than direct evidence.

Conceptual and secondary publications contributed mainly to the interpretation of opportunities and risks. Kasneci et al. (2023) identified personalization, immediate feedback, content generation, and learner support as potential benefits of generative AI, while also emphasizing inaccuracies, bias, dependence, privacy concerns, and academic integrity. Lo (2023) reached similar conclusions regarding educational assistance and personalized explanations but highlighted concerns about assessment and responsible use. Farrokhnia et al. (2024) emphasized both the strengths of immediate interaction and personalized responses and the threats associated with misinformation, excessive dependence, and academic dishonesty. Bond et al. (2024) further argued that AI adoption in higher education requires



ethical principles, collaboration, rigorous research, and evidence-based institutional policies. Some findings were coincident, whereas others were less consistent. Most empirical studies agreed that generative AI can support preparation, feedback, interaction, and student participation (Huesca et al., 2024; Yavuz et al., 2025; Cabrera-Solano et al., 2026; H. Liu et al., 2026). Nevertheless, the evidence was more variable regarding academic achievement and deeper learning. Huesca et al. (2024) and H. Liu et al. (2026) reported improvements in academic performance, while Y. J. Liu et al. (2026) found that the AI-enhanced model was not superior in all learning areas. These apparently contradictory results may be explained by differences in discipline, intervention duration, AI tool, instructional design, assessment procedure, teacher mediation, and student preparation.

Disciplinary context may also influence the observed outcomes. In programming and language education, generative AI can provide examples, corrections, explanations, and immediate feedback that directly support practice and revision (Apaza Calsin et al., 2023; Huesca et al., 2024; Cabrera-Solano et al., 2026). In mental health and pharmacology education, however, the consequences of inaccurate or insufficiently contextualized information may be more significant, increasing the importance of teacher supervision and verification (H. Liu et al., 2026; Y. J. Liu et al., 2026). Therefore, the effectiveness of AI-supported flipped learning cannot be generalized without considering the complexity and specific requirements of each discipline.

The methodological limitations of the included evidence must also be considered. The corpus combined empirical studies, experimental and quasi-experimental research, systematic reviews, meta-analyses, rapid reviews, and conceptual publications. These designs differed in participants, intervention duration, AI tools, disciplines, outcome measures, and analytical procedures, which prevented direct comparison and statistical synthesis. In addition, several empirical studies were conducted in specific courses or institutional contexts, limiting the generalizability of their results. The inclusion of secondary and conceptual evidence also creates a risk of giving similar interpretive weight to publications that do not provide direct evidence of student outcomes.

The review process itself presents additional limitations. Study identification, selection, data extraction, thematic analysis, and quality appraisal were performed by one reviewer, and no inter-rater agreement coefficient was calculated. The review was not prospectively



registered, and no public protocol was published before the search. Consequently, selection bias, subjective interpretation, or unintentional omissions cannot be excluded. The descriptive quality matrix also assessed reporting characteristics rather than formal risk of bias, and its scores should not be interpreted as proof that all included studies had equivalent methodological rigor.

Overall, generative AI appears most educationally valuable when it functions as a guided learning assistant rather than as an automatic answer generator. The reviewed evidence suggests that its contribution depends on structured pre-class preparation, active classroom tasks, critical verification, teacher mediation, and explicit ethical guidelines (Lee et al., 2024; Kasneci et al., 2023; Farrokhnia et al., 2024; Bond et al., 2024). Future research should conduct longer interventions, use larger and more diverse samples, compare conventional and AI-supported flipped models, and report separately the effects on academic performance, participation, self-regulation, higher-order thinking, and emotional well-being.

5. CONCLUSIONS

This integrative review analyzed the contribution of generative artificial intelligence-supported flipped classrooms in higher education. The literature search covered the period from 2020 to 2026, and 15 publications were included in the final corpus. The evidence was organized into four main dimensions: academic achievement and learning performance; student engagement and active participation; self-regulated learning and higher-order thinking; and pedagogical, technological, and ethical implementation conditions.

The reviewed evidence indicates that generative AI can strengthen flipped learning by supporting pre-class preparation, personalized explanations, immediate feedback, active classroom participation, self-regulation, and higher-order thinking. The most consistent benefits appeared when students used AI to prepare, formulate questions, evaluate information, and apply knowledge during structured classroom activities.

However, the educational value of AI depended on instructional design, teacher guidance, technological access, student responsibility, and ethical use. The evidence did not show that AI-supported flipped classrooms produce identical benefits across all disciplines, students, or learning outcomes. In some cases, AI improved preparation and efficiency without producing stronger results in deeper disciplinary learning.



The findings also showed that generative AI should not be used as an automatic answer generator or as a replacement for teachers. Its most appropriate role is to function as a learning assistant that provides explanations, guidance, feedback, and opportunities for reflection. Students should remain responsible for verifying information, explaining their reasoning, and constructing their own conclusions.

The conclusions are limited by the methodological diversity of the 15 publications, the inclusion of both primary and secondary evidence, the participation of one reviewer, and the absence of prospective protocol registration. Future research should prioritize longer and more comparable interventions, larger and more diverse samples, transparent reporting, and direct comparisons between conventional and AI-supported flipped classrooms.

Generative artificial intelligence-supported flipped learning should therefore be understood as a structured pedagogical model rather than simply the use of ChatGPT or other AI tools before class. When responsibly designed, it can strengthen higher education without replacing teacher mediation, active learning, critical thinking, or student responsibility.

REFERENCES

- Apaza Calsin, M., Aedo, M., & Castro, E. (2023). Impacto de ChatGPT en la enseñanza: Un enfoque de aula invertida para fundamentos de programación. *RISTI–Revista Ibérica de Sistemas e Tecnologias de Informação*, 52, 97–112. <https://doi.org/10.17013/risti.52.97112>
- Baig, M. I., & Yadegaridehkordi, E. (2023). Flipped classroom in higher education: A systematic literature review and research challenges. *International Journal of Educational Technology in Higher Education*, 20, Article 61. <https://doi.org/10.1186/s41239-023-00430-5>
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21, Article 4. <https://doi.org/10.1186/s41239-023-00436-z>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in reflexive thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>



- Bredow, C. A., Roehling, P. V., Knorp, A. J., & Sweet, A. M. (2021). To flip or not to flip? A meta-analysis of the efficacy of flipped learning in higher education. *Review of Educational Research*, 91(6), 878–918. <https://doi.org/10.3102/00346543211019122>
- Cabrera-Solano, P., Castillo-Cuesta, L., & Ochoa-Cueva, C. (2026). Active learning and feedback in EFL teacher education through AI-supported flipped classrooms. *Education Sciences*, 16(6), Article 827. <https://doi.org/10.3390/educsci16060827>
- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, 6, Article 100198. <https://doi.org/10.1016/j.caeai.2023.100198>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. <https://doi.org/10.1080/14703297.2023.2195846>
- Huesca, G., Martínez-Treviño, Y., Molina-Espinosa, J. M., Sanromán-Calleros, A. R., Martínez-Román, R., Cendejas-Castro, E. A., & Bustos, R. (2024). Effectiveness of using ChatGPT as a tool to strengthen benefits of the flipped learning strategy. *Education Sciences*, 14(6), Article 660. <https://doi.org/10.3390/educsci14060660>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21, Article 16. <https://doi.org/10.1186/s41239-024-00447-4>
- Liu, H., Sriwisathiyakun, K., & Petsangsri, S. (2026). Integrating generative AI with the flipped classroom teaching model: Enhancing emotional well-being and student



achievement. *Cogent Education*, 13(1), Article 2658290.
<https://doi.org/10.1080/2331186X.2026.2658290>

Liu, Y. J., Ma, H. Z., Luo, H. Y., Xiong, Y. X., Qu, F. W., Xie, J. P., & Guo, Y. (2026). From traditional classroom to AI-enhanced flipped classroom: A three-year pedagogical evolution for international students in pharmacology. *BMC Medical Education*, 26, Article 1111. <https://doi.org/10.1186/s12909-026-09608-7>

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), Article 410.
<https://doi.org/10.3390/educsci13040410>

Lubbe, W., ten Ham-Baloyi, W., & Smit, K. (2020). The integrative literature review as a research method: A demonstration review of research on neurodevelopmental supportive care in preterm infants. *Journal of Neonatal Nursing*, 26(6), 308–315.
<https://doi.org/10.1016/j.jnn.2020.04.006>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71.
<https://doi.org/10.1136/bmj.n71>

Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review*, 30, Article 100314. <https://doi.org/10.1016/j.edurev.2020.100314>

Yavuz, M., Balat, Ş., & Kayalı, B. (2025). The effects of artificial intelligence-supported flipped classroom applications on learning experience, perception, and artificial intelligence literacy in higher education. *Open Praxis*, 17(2), 286–304.
<https://doi.org/10.55982/openpraxis.17.2.811>

CONFLICTS OF INTEREST

None.

FUNDING

Not applicable.



AUTHOR CONTRIBUTIONS (CRediT)

Marlene Elizabeth Rengifo Obando: conceptualización, investigación, metodología, redacción del borrador original y revisión y edición del manuscrito.

Juan Gabriel Vera Alcivar: investigación, recopilación de información, análisis formal, redacción y revisión del manuscrito.

Carolina de Lourdes Arcentales Hidalgo: investigación, curación de datos, organización de la información, redacción y edición del manuscrito.

Karen Sabrina Brito Ordoñez: investigación, visualización, recopilación de información y revisión y edición del manuscrito.

Poleth Alejandra Mendieta Sinche: investigación, recopilación de información, curación de datos, redacción y revisión y edición del manuscrito.

