South American Publishing

SAP Gamification and Augmented Reality. 2025;4:270
doi: 10.62486/gr2025270

REVIEW

Generative artificial intelligence as game designer and tutor in Latin American classrooms

Inteligencia artificial generativa como diseñadora de juegos y tutora en aulas latinoamericanas

Analía Ortigoza[1]
Lucía Gerez[1]

[1] Universidad Nacional de Tucumán. Tucumán, Argentina

Submitted: 30-06-2025 | Revised: 12-09-2025 | Accepted: 04-11-2025 | Published: 07-11-2025

Corresponding Author: Analía Ortigoza

ABSTRACT

Generative artificial intelligence is entering Latin American classrooms as a tutor and as an aid to game design, but the regional evidence has not been synthesized by role. This narrative review summarizes studies set in or about Latin America and published from 2013 to 2025. Fifty-six works were selected. Sixteen core works addressed tutoring, 14 addressed gamified or game-based activities with artificial intelligence, 10 addressed risks, 10 addressed adoption and teacher readiness, and six were reviews. Tutor studies reported faster task completion, feedback at scale and mixed learning gains, including two randomized trials with opposite results. Generative models appeared as game designers mainly in teacher preparation of playful activities, and none of the gamification works compared them with conventionally designed games. The reported risks were plagiarism, lower critical thinking with dependence, inaccurate output, and unequal access by region and gender. Only abstracts were read, and the coding awaits verification by the authors. Designs that separate the generative component, scaffold the tutor and train users in verification are the most direct way to test these claims.

Keywords: generative artificial intelligence; educational games; intelligent tutoring systems; gamification; Latin America.

RESUMEN

La inteligencia artificial generativa llega a las aulas latinoamericanas como tutora y como apoyo al diseño de juegos, pero la evidencia regional no se ha sintetizado por rol. Esta revisión narrativa resume estudios ubicados en América Latina o sobre ella y publicados entre 2013 y 2025. Se seleccionaron 56 trabajos. Dieciséis trabajos centrales trataron la tutoría, 14 las actividades gamificadas o basadas en juegos con inteligencia artificial, 10 los riesgos, 10 la adopción y la preparación docente y seis fueron revisiones. Los estudios sobre tutoría informaron mayor rapidez en las tareas, retroalimentación a escala y ganancias de aprendizaje mixtas, con dos ensayos aleatorizados de resultado opuesto. Los modelos generativos aparecieron como diseñadores de juegos principalmente en la preparación docente de actividades lúdicas, y ningún trabajo de gamificación los comparó con juegos diseñados de forma convencional. Los riesgos descritos fueron el plagio, el menor pensamiento crítico con la dependencia, la salida inexacta y el acceso desigual según región y género. Solo se leyeron resúmenes y la codificación espera la verificación de los autores. Los diseños que separan el componente generativo, andamian la tutoría y entrenan en verificación son la vía más directa para probar estas afirmaciones.

Palabras clave: inteligencia artificial generativa; juegos educativos; sistemas de tutoría inteligente; gamificación; América Latina.

© 2025; The Authors. This Open access article is distributed under CC BY 4.0 licence.

INTRODUCTION

Chatbot tutors raised learning outcomes in a meta-analysis of 24 randomized studies, with larger effects in higher education than in primary or secondary schooling.(1) A later synthesis of 62 studies and 135 effect sizes reported a large effect that fell to small or moderate after correction for publication bias.(2) Large language models add content creation and personalized interaction on demand, and they require fact-checking and critical judgment from the user.(3) Their tendency to hallucinate, that is, to produce plausible but nonfactual text, is documented.(4) Dialogue-based tutors predate these models: a review of 33 of them found evaluations of learning gains for 90.91% of the systems.(5)

Regional evidence is growing faster than its synthesis. A systematic review of 25 papers published from 2021 to 2023 covered generative AI in higher education across eight Latin American countries,(6) and a scoping review of 84 documents from 2022 to 2024 mapped the competencies teachers need.(7) Student data can be large: 4,811 students at eight Ecuadorian universities answered one plagiarism questionnaire.(8) Structural inequality conditions any deployment, as a review of 28 articles linking the gender digital divide to education in Latin American countries indicated.(9)

Gamification adds a second thread. A meta-analysis reported significant effects on cognitive ($g = 0.49$), motivational ($g = 0.36$) and behavioral ($g = 0.25$) learning outcomes,(10) and a later meta-analysis of 22 experimental studies found a moderately positive effect on academic performance ($g = 0.782$).(11) Generative models now reach the design step itself. ChatGPT was proposed for brainstorming themes and mechanisms, drafting game components and critiquing prototypes in serious board game design.(12) A language-model tool generated gamified programming exercises, most of which were ready for use with little human effort.(13)

Designers are less enthusiastic than tool builders. In a survey of 42 game design professionals, generative AI accelerated ideation and prototyping, yet respondents raised concerns about originality and creative dependency.(14)A narrative review of the game industry named playability, narrative originality, bias, intellectual property and content homogenization as persistent challenges.(15)None of these sources was set in a Latin American classroom, and the question of what the region has documented for classrooms remains open.

The objective of this narrative review was to synthesize studies set in or about Latin America on generative AI in two classroom roles, tutor and designer of games and gamified activities, and to record the opportunities, risks and conditions of adoption they report. It asks which tutor roles are documented, how generative models take part in designing games and gamified activities, and which risks recur. Its scope differs from earlier contributions to this journal on a prompt engineering game and on ChatGPT in libraries: it concerns teaching and learning in schools and universities, and it neither treats prompting as a game nor addresses library services.

DEVELOPMENT

The 56 core works came from Ecuador (n = 17), Peru (n = 14), Brazil (n = 5), Colombia (n = 5), Chile (n = 3), Mexico (n = 2), Uruguay (n = 1) and Guatemala (n = 1), plus two multi-country studies and six reviews. Sixteen works addressed tutoring, 14 addressed games and gamified activities, 10 addressed risks, and 10 addressed adoption and teacher readiness. Seventeen titles were in Spanish or Portuguese. Only two of the 16 tutor works were randomized, and four others were quasi-experimental (Table 1).

The strongest tutor design was a randomized trial of 162 diplomacy students in Colombia, in which ChatGPT-assisted debates produced higher combined scores (multivariate analysis of covariance) and gains in understanding of complex concepts, critical thinking and argumentation.(16) Instead, a randomized experiment with 41 master's students in Ecuador found no significant gain in digital storytelling skills, although germane cognitive load fell in the ChatGPT group.(17)The two randomized comparisons therefore pointed in different directions.

Programming courses supplied two comparisons. In Chile, 71 industrial engineering students learned web programming with Microsoft Copilot or with instructional videos; both groups improved, and the video group gained more knowledge and showed more curiosity and stronger intention to continue.(18)At the University of Brasilia, 41 students in an introductory course worked with human teaching assistants or with ChatGPT for three weeks. The ChatGPT group completed more assignments in less time with comparable accuracy, yet many students still preferred human help.(19)

Table 1. Synthesis of the 56 core works by axis.

Axis	Works (n)	Countries	Designs	Main reported result	Main limit
Tutor and assistant	16	Peru 7, Chile 3, Ecuador 2, Mexico 2, Brazil 1, Colombia 1	2 randomized, 4 quasi-experimental, 1 pre-experimental, 9 other	Faster task completion and feedback; mixed learning gains	Samples of 22 to 347 where reported; short interventions
Game and gamified design	14	Ecuador 9, Brazil 3, Colombia 2	4 quasi-experimental, 1 action research, 1 game-data analysis, 1 rapid review, 6 surveys, 1 qualitative	Teacher-side preparation of playful activities; motivation reported	AI and gamification confounded; no comparison with conventional games
Risk	10	Peru 3, Ecuador 3, Brazil 1, Colombia 1, Uruguay 1, three-country 1	6 cross-sectional, 3 qualitative, 1 framework with a planned RCT	Plagiarism, dependence, unequal access	Self-report; samples of 9 to 4,811
Adoption and teacher readiness	10	Peru 4, Ecuador 3, Colombia 1, Guatemala 1, five-country 1	1 quasi-experimental, 8 surveys or mixed methods, 1 qualitative	Adoption follows perceived benefit and peers; teacher training gaps	Perceptions, not learning outcomes
Review	6	Latin America	3 systematic, 1 scoping, 2 literature reviews	Opportunities with ethical and infrastructure limits	Overlap with core studies not assessed

Note: RCT = randomized controlled trial.

Two Peruvian designs tested how students use the tool. A prompting method (role, objective, context, examples and format) taught to 30 of 60 university students raised critical thinking scores, with a median of 10 against 8 in the appropriability dimension ($d = 2.042$).(20)In a quasi-experiment with 100 nursing students in Trujillo, a 15-session structured use of ChatGPT improved research skills against a comparison group ($U = 95.4$, $p < 0.05$).(21)Effect sizes this large in samples of 60 and 100 call for caution.

Two comparisons qualify the gains. In a Mexican course, 96 students formed two groups; the ChatGPT group obtained better grades on practical than on theoretical tasks, whereas the group without ChatGPT did better in a final exam taken without technological assistance. The ChatGPT group reported higher satisfaction and motivation.(22)In a Peruvian software engineering program, ChatGPT improved task efficiency and gave real-time help, yet students admitted relying on it for analytical tasks.(23)

Tutor prototypes were small. A 2 × 2 experiment with 22 engineering students in Mexico crossed voice or text interaction with the presence of an avatar: voice lengthened active time, text produced more messages, and the avatar changed none of the measured variables.(24) A chatbot for secondary students in Lima was described through its cloud architecture and a planned pilot, so no learning outcome was available.(25) In Peru, 53 first-year students who used a conversational agent improved in the social dimension of adaptability ($p < 0.001$) but not in the institutional one ($p = 0.124$).(26)

An intelligent tutoring study of systems engineering students at a Peruvian university reported R^2 values of 0.746 for student emotion and 0.723 for the effect on academic performance.(27)A tangible programming method in Lima used a multimodal language model to translate photographed cards into robot code. In a pilot of 60 pupils, 94% of the robot group scored 5 of 5 against 65% of the unplugged group.(28)The model acted as translator, not as tutor.

Feedback on written work was a second use. In Chile, ChatGPT-4 with human validation (a process named RECaP-GPT) produced descriptive feedback on 578 responses to pedagogical cases written by 347 students in seven teacher education programs. The feedback followed criteria known to the students, and the authors listed implementation limits.(29)

Accuracy and prompting set the ceiling of any tutor. ChatGPT scored 8.25 against 7.35 for 110 medical students on four clinical cases in Ecuador ($p = 0.0293$), and the authors warned about inaccurate information.(30) Among 102 Chilean students writing Spanish prompts, explicit context raised response quality more than stating the purpose alone.(31) For English speaking, a review of Latin American work from 2024 and 2025 described ChatGPT as a supplementary tool for overpopulated classrooms.(32)

Three patterns recur in the tutor axis. Speed and completion improved more consistently than unaided learning, as the Brasilia and Mexican courses show. The largest reported effects came from structured interventions, the prompting method and the 15-session nursing program, with groups of 60 and 100 students. School-level evidence was limited to the Lima chatbot design and

the card-based pilot, and the other 14 tutor works involved university students.

None of the 14 game works compared a game built with a generative model against a conventional game. The closest was a workshop in which 24 prospective physics teachers in Ecuador used ChatGPT to design playful activities. Of them, 79% valued its adaptation of activities to learner levels and 83% its reduction of preparation time.(33)

An eight-week quasi-experiment with 60 early-childhood teacher trainees in Ecuador gave the experimental group (n = 30) adaptive AI feedback, educational chatbots, badges, levels and H5P mini-games created with an AI tutor. More than 70% of them increased voluntary participation, sustained attention and collaboration.(34)The design confounds gamification and AI, so the generative component cannot be isolated.

Pupils with dyslexia or language disorders (n = 80) who used AI-driven game platforms gained in vocabulary, reading and writing, and overall performance rose 27% after the intervention, while the control group received traditional instruction.(35)Teacher surveys outnumbered trials. In Quevedo, 48 teachers perceived that gamification combined with generative AI favored problem solving, creativity, personalized learning and critical thinking.(36)

A mixed-methods study of 40 public institutions in Ecuador, with 400 teachers and 800 students, found high teacher acceptance of AI and student motivation toward gamified activities. Technology gaps were most severe in rural areas and the Amazon.(37)An action-research design in a fourth-grade natural sciences class reported relevance and effectiveness after validation of gamification and AI strategies.(38)

In Brazil, a business budgeting game for accounting students used k-nearest neighbors, random forest and artificial neural networks as opponents, and four semesters of student decisions were compared with those of the algorithms.(39) These are predictive models, not generative ones. A rapid review of seven Portuguese-language articles found that AI could optimize question creation in quiz games and sustain support after assessment through chatbots.(40) Qualitative evidence on barriers came from a systematic review and a focus group in Ecuador: resistance to change, administrative workload, training needs and institutional limits hindered gamification, while AI was seen as a way to build didactic resources in a short time.(41) In Colombia, 78% of 50 accounting students rated AI as highly relevant and 66% valued gamification, and the barriers were lack of training and resistance to change.(42)

Teacher-design work was the closest regional match to the designer role. A training plan in an Ecuadorian school combined ChatGPT for pedagogical text, Curipod for real-time interactive presentations and MagicSchool.ai for differentiated materials and formative assessment. After seven sessions, a quasi-experimental evaluation showed gains in technical appropriation, interactive design capacity, evaluative use of the platforms and willingness to adopt them.(43)

In secondary language and literature classes, a quasi-experiment with control and experimental groups combined project-based learning, universal design for learning, gamification and AI. It reported better intrinsic motivation, reading comprehension and text production, and a 27% rise in communicative performance over traditional techniques.(44)

Teachers used ChatGPT as a content source. Among 86 Brazilian computer science professors, it served to generate exercises, slides, assignments and tests.(45)EdutekaLab, a platform of Universidad Icesi in Colombia that includes a gamification tool, was assessed by 408 Latin American educators, who reported benefits for creativity, critical thinking and collaboration.(46)Only three of the 14 game works reported a control group.

The game works split by who designed the activity. Teachers used AI to prepare activities in four works,(33,41,43,45)whereas pupils or trainees met AI-supported games in the others. Only four of the 14 abstracts named ChatGPT or generative AI explicitly, and one work used classical machine learning. The generative designer role is therefore inferred, not tested.

Adoption studies explained why the tools spread. Among 792 undergraduates in Peru, Chile, Bolivia, Argentina and Colombia, relative advantage (β = 0.247) and compatibility (β = 0.246) had the largest effects on attitude toward ChatGPT, and attitude predicted intention to use (β = 0.777).(47)Adoption followed perceived academic benefit more than technical ease. Peer influence also mattered: among 562 Peruvian students, word of mouth (β = 0.569) and social norms (β = 0.253) predicted intention to use chatbots, and intention predicted actual use (β = 0.711).(48)

Use varied by group. Among 256 students from six Ecuadorian universities, female and engineering students showed higher adoption, and four user clusters emerged.(49)In Guatemala, 260 students used chatbots chiefly to prepare class presentations (68.4%) and to humanize texts (66.7%).(50)

Teachers lagged students in some settings. At the University of Córdoba in Colombia, students showed high familiarity and frequent use, while faculty adopted generative AI more cautiously.(51)

Among 180 Peruvian instructors, higher TPACK levels were moderately linked to more frequent and purposeful use of ChatGPT, and gender, age and prior AI training also mattered.(52)

Ten writing teachers at a private university in Lima used ChatGPT mainly for the final revision of texts and little for planning, in a rather unreflective manner.(53)A quasi-experiment with 60 Peruvian lecturers (30 per group) tested 14 online workshops over eight weeks. Digital competence rose in all five dimensions ($p < 0.001$; $d = 2.05$).(54)

Perceived benefits were similar in language classes. Among 120 EFL students in Ecuador, ChatGPT was seen as helpful for organizing ideas, managing time and improving writing, with concerns about overdependence, lack of personalization and occasional inaccuracies.(55)In Ecuadorian secondary education, real-time individualized explanations were reported to improve motivation, task dedication and critical thinking, alongside concerns about teacher preparation, dependence and accuracy.(56)

Two literature reviews addressed school levels. For secondary education, generative tools served mainly to search for information, draft texts, organize ideas and obtain feedback, and substitution was linked to less verification and more passivity.(57)In Brazilian basic education, a systematic review of six empirical studies found that teachers used generative AI for lesson plans, examples, assessments and inclusive materials, and met conceptual inaccuracies, curricular misalignment, weak infrastructure and limited digital literacy.(58)

At eight Ecuadorian universities, a logistic model explained 20.2% of the variance in plagiarism level (Nagelkerke R^2), and trust in ChatGPT-generated information (OR = 1.176) and reporting its use to instructors (OR = 1.223) were associated with higher plagiarism levels. Knowing more AI tools (OR = 0.911) and perceiving that AI improves performance (OR = 0.945) were associated with lower levels.(8)In Colombia, 46.7% of 381 students used ChatGPT for evaluative assignments, 60.4% let it complete parts of assignments, 17.1% used it to complete at least one of their last four assignments in full, and 7.6% had received training.(59)

Among 56 Ecuadorian EFL students, worry about hindering their own writing development exceeded worry about penalties. Submitting chatbot text was judged dishonest, and fewer participants judged machine translation from Spanish dishonest.(60)

Use of generative AI was moderate and mostly supportive in 200 first-semester students in Trujillo, yet dependence was negatively associated with overall critical thinking, more strongly for inference and evaluation.(61) In Lima, more than 250 STEM students showed no academic benefit from frequent AI use, and some showed lower performance, whereas higher technological self-efficacy went with better results.(62) In turn, among 833 Ecuadorian students, emotional ($\beta = 0.413$) and cognitive ($\beta = 0.567$) dimensions predicted ethical behavior, while frequency of use did not ($\beta = -0.128$; $p = 0.058$).(63)

In the Ceibal program in Uruguay, female teachers and students engaged more with preparation and structured practice tools, and male counterparts dominated exploratory platforms. The authors designed a randomized trial of a chatbot structured for teacher mediation.(64) Qualitative work in Peru recorded concerns about dehumanization, the digital divide and inequity in access.(65)

Nine upper-secondary teachers in Brazil saw gains in teaching quality and student interest, and named plagiarism and weak digital competence as obstacles.(66)Interviews with 26 decision-makers and staff in Ecuador, Chile and Costa Rica located difficulties in discerning human from AI-generated content, in promoting critical thinking and in student inequalities.(67)

Risk works differ in what they measured: integrity in three, cognition in two, ethical behavior in one, and equity or teacher and institutional constraints in four. Nine of the 10 relied on self-report or interviews, and the Uruguayan study used platform data and classroom observation. None of them tracked the same students over time.

Adoption works converged on perceived benefit, peers and training. Eight of the 10 used surveys or mixed methods, and only one, the lecturer workshops, measured the result of a training. Only one adoption study was set in schools, in Ecuadorian secondary education, and the remaining nine involved universities.

The six reviews gave the widest view. A systematic review of 25 papers covered applications, perceptions and institutional challenges across eight countries.(6)A PRISMA-based review identified opportunities in computer science and engineering alongside ethical limits.(68)A scoping review of 84 documents described five dimensions of teacher competence, among them creative adaptation to limited infrastructure.(7)The remaining three reviews, on English speaking, secondary education and Brazilian basic education, appear above.

Across axes, three findings cut through the regional corpus. Tools were adopted before they were evaluated, since surveys outnumbered comparative designs in the game, risk and adoption axes. Gains were reported mostly for completion, motivation and perceived usefulness, and in few

studies for unaided learning. Risks were described by the same populations that used the tools, which gives their self-reports weight but no external check.

DISCUSSION

Regional evidence casts generative AI mainly as tutor and teaching assistant, and much less as game designer. Tutor works were more numerous (16 against 14) and more often experimental, whereas the game axis rested on teacher surveys and bundled interventions. The designer role appears in Latin American classrooms as a teacher task, such as preparing playful activities, interactive materials or exercises,(33,43,45)and not as an evaluated, student-facing capability.

Opportunities cluster in saved time and in feedback, risks cluster in integrity, dependence and unequal access, and adoption follows perceived benefit.(47)This pattern is consistent with the position that language models help create content and personalize interaction but demand fact-checking and critical thinking.(3)The regional record adds data on how often students hand over assignments and how rarely they are trained.(59)

Concordant evidence on tutoring comes from outside the region. In a randomized trial, students learned significantly more in less time with a custom AI tutor than in an active learning class,(69)which parallels the Brazilian finding of more assignments completed in less time(19)and the Colombian gains in a randomized debate design.(16)Chatbot meta-analyses also report positive effects,(1,2) and a broader review of AI learning tools in higher education linked them to performance and engagement gains.(70) For schools, where regional tutor evidence is scarcest, a systematic review of 20 studies with 2,853 K-12 students found generally positive effects of intelligent tutoring, mitigated when the comparison was a non-intelligent tutoring system.(71)

Concordant findings on feedback and dependence are also available. A review of generative chat feedback reported better text structure and cohesion,(72)in line with the Chilean feedback process.(29)In a randomized experiment, learners supported by ChatGPT improved essay scores, but knowledge gain and transfer did not differ significantly from other groups.(73)This matches the null storytelling result in Ecuador(17)and the negative association between dependence and critical thinking in Peru.(61)

Mediation emerges as a shared condition. A review on secondary education concluded that the impact depended more on interaction quality, pedagogical purpose and teacher mediation than on frequency of use,(57)and the Ecuadorian ethics model likewise found that frequency alone did not predict ethical behavior.(63)Teacher-side data agree: Brazilian professors generated exercises and tests,(45)Peruvian writing teachers used the tool mostly for revision,(53)and Colombian faculty adopted it more cautiously than students.(51)

On games, concordance is partial. Gamification meta-analyses report positive effects,(10,11) and the Ecuadorian trials of gamified AI interventions also reported gains.(34,35) Generative tools were judged ready for use as exercise generators(13)and as brainstorming partners,(12)and designers reported faster ideation.(14)These sources agree with the 83% of Ecuadorian trainees who reported shorter preparation time.(33)

Several results diverge. In Chile, instructional videos outperformed Copilot in early programming, whereas a custom tutor built on pedagogical best practices outperformed active learning outside the region.(18,69) A plausible explanation is design: a general-purpose assistant answers, and a scaffolded tutor structures learning. Chatbot meta-analyses disagree on duration: one found short interventions stronger and attributed this to novelty,(1)and another found longer interventions more effective.(2)The regional trials lasted from three to eight weeks, so novelty may have inflated their effects, a hypothesis the studies did not test.

Avatars diverge too. A meta-analysis recommended human-like avatars for chatbots,(1)yet the Mexican experiment found that an avatar altered no measured variable.(24)An LLM tutor inside an educational game for language learners produced only marginally significant accuracy gains and signs of over-reliance,(74)a more cautious result than the Ecuadorian mini-game reports. After bias correction, the chatbot effect shrank to small or moderate.(2)

Short-term performance and unaided learning also diverge. The Mexican group that used ChatGPT scored higher on practical tasks but lower on an unaided final exam,(22)and the Peruvian engineering students admitted relying on the tool for analytical tasks.(23)Both observations echo the essay-score gain without transfer in the randomized experiment,(73)and they suggest that task completion is a poor proxy for learning.

Integrity findings diverge in what predicts misconduct. In Ecuador, broader knowledge of AI tools was associated with less plagiarism,(8)and ethical behavior depended on reflection and

knowledge, not on frequency of use.(63)This suggests, without establishing causality, that training may matter more than restriction. Detection is a weak substitute, since GPT detectors misclassified non-native English writing as AI generated,(75)an inference that bears on Latin American learners of English but was not tested in the region.

Where numbers exist, the opportunities are specific. Of the Ecuadorian trainees, 83% reported shorter preparation time;(33)three weeks of ChatGPT assistance in Brasilia allowed more assignments at comparable accuracy;(19)a 27% rise in language performance was reported twice in Ecuador;(35,44) and 578 student responses received formative feedback in Chile.(29)Each figure comes from a single study, and the corpus contains no replication of any of them.

Four mechanisms are plausible, and each has partial support. Reduced mental effort can lower germane cognitive load without raising skill, as in the Ecuadorian storytelling trial.(17)A randomized experiment on the risk its authors called metacognitive laziness found essay score gains without differences in knowledge gain and transfer.(73)Technological self-efficacy can moderate the benefit, since Lima students with higher self-efficacy performed better whatever their AI use.(62)Structure supplies the fourth, because scaffolded tutoring and trained prompting accompanied the clearest gains.(20,69) None of the mechanisms was tested directly in the region.

Accuracy remains a design constraint. Hallucination is a documented property of language models,(4) and the Ecuadorian clinical case test showed higher scores for ChatGPT than for students while the authors warned of inaccuracies.(30)The case for tutors in overpopulated classrooms(32) depends on that accuracy, and no regional study measured it in a live classroom.

Interaction design bears on the usability concerns of the journal. A meta-analysis found that text-based interactions showed the largest chatbot effects,(2)while the Mexican tutor experiment found that voice lengthened active time and text produced more messages.(24) Satisfaction can also mislead: a game-embedded tutor received positive ratings with limited learning gains,(74)and the Mexican ChatGPT group reported higher satisfaction yet scored lower on the unaided exam.(22)

Language and context matter for tutors. Spanish prompts with explicit context improved response quality,(31)Ecuadorian EFL students valued ChatGPT for writing but feared overdependence,(55)and they worried more about their own skills than about penalties.(60)Detectors that misclassify non-native English writing(75)could penalize precisely these learners. Regional tutor evaluations therefore need Spanish and Portuguese protocols, not only translations of English ones.

Equity conditions every finding. A review linked the gender digital divide to education across Latin American countries,(9) gender patterns of tool use appeared in Uruguay,(64)and female and engineering students adopted ChatGPT more in Ecuador.(49)Rural and Amazonian access gaps(37) and weak infrastructure in Brazilian basic education(58)limit who can benefit. Estimates from university samples may therefore overstate what public rural schools can expect.

Game-design evidence from outside the region defines what Latin American trials should measure. Playability, narrative originality, bias and content homogenization are the stated challenges,(15)designers worry about originality and creative dependency,(14)and an LLM tutor inside a game produced limited learning gains despite positive satisfaction.(74)None of these outcomes was reported in the abstracts of the 14 regional game works.

Earlier regional reviews covered AI in higher education,(6)teacher competencies,(7)opportunities and ethical limits(68)and English speaking.(32)Their conclusions on ethical limits, infrastructure and teacher training agree with the present synthesis. This review adds a separation by classroom role, school-level sources, and the game axis, where the regional record is thinnest.

Strengths are limited but real. Queries ran in three languages, the regional coverage is declared by country, and each statement rests on the abstract of a verifiable record with a stable identifier. Counts of designs follow the wording of abstracts, so no study was upgraded to a randomized or controlled design that its abstract did not state.

Limitations are concrete. Only abstracts were read, so design details and effect sizes were not verified in full texts. Double coding of full texts with a formal risk-of-bias tool, such as RoB 2 for randomized trials, would remove this limit.

Timing limits the evidence. Of the 56 core works, 47 were published in 2025, eight in 2024 and one in 2023, so most observations followed the arrival of ChatGPT by a single academic cycle. Cohort follow-up over a full school year, with the same instruments at entry and exit, would show whether novelty effects fade.

Coverage is also uneven. Ecuador and Peru contributed 31 of the 50 primary studies, Argentina and Bolivia appeared only inside one five-country survey, and SciELO, Redalyc and LILACS were not searched directly. Direct searches in those indexes, with the records found reported, would correct the imbalance.

Most game works bundled gamification and AI, only three reported a control group, one

used classical algorithms, and 13 of the 14 appeared in journals not matched to the Scopus catalogue used for this review. Some effect sizes were very large in small groups, such as d = 2.042 with 60 students and d = 2.05 with 60 lecturers. Factorial designs that vary the generative component and the game elements separately would isolate each effect, and prespecified outcomes would temper the small-sample estimates.

Screening and coding with an AI assistant also await independent human verification of the matrix. A second coder who rebuilds the counts of Table 1 from the abstracts, and who resolves disagreements by reading the full texts, would settle whether the axis assignments are reproducible.

Practical implications follow. Tutors should be scaffolded and tested against active teaching.(69)Students need training in verification, because only 7.6% of the surveyed Colombian students had it,(59)and the Peruvian lecturer workshops show that short training can raise digital competence.(54)Teacher mediation and attention to gender patterns of use are design requirements.(64)Rural and Amazonian access gaps call for infrastructure policy before adoption,(37)and ethical frameworks must accompany deployment.(76)

For games, generative drafts of activities should pass a human playtest for playability, originality and bias.(15)Trials should evaluate games that teachers or pupils build with these models against conventionally designed games, with usability and engagement measured. Assessment must change with the tools. Where 60.4% of surveyed students let ChatGPT complete parts of assignments,(59)take-home tasks measure assistance more than learning. Unaided exams, oral defenses and process evidence give a cleaner signal, and the Mexican contrast between practical tasks and an unaided exam shows why.(22)Detection software cannot carry that burden.(75)A feasible regional agenda has three designs. The first is a randomized comparison of a scaffolded tutor against a general-purpose assistant, with unaided post-tests, as the Chilean and Mexican contrasts suggest. The second is a factorial trial that crosses game elements with a generative component. The third is a design study in which teachers and pupils build games with a language model, followed by playtests and measures of usability and engagement. That agenda fits the human-computer interaction and game design concerns of the journal.

CONCLUSIONS

Regional studies place generative AI as a tutor that speeds task completion and feedback, with learning gains that depend on design and that were absent in some randomized comparisons. As a game designer, it appears chiefly in teacher preparation of playful activities, and none of the retrieved works compared it with conventional game design. Plagiarism, dependence and unequal access were the recurring risks, and they call for scaffolded designs, training in verification and trials that separate the generative component from gamification.

REFERENCES

1. Wu R, Yu Z. Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. Brit J Educational Tech. 2024;55(1):10-33. doi: 10.1111/bjet.13334
2. Laun M, Wolff F. Chatbots in education: Hype or help? A meta-analysis. Learning and Individual Differences. 2025;119:102646. doi: 10.1016/j.lindif.2025.102646
3. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences. 2023;103:102274. doi: 10.1016/j.lindif.2023.102274
4. Huang L, Yu W, Ma W, Zhong W, Feng Z, Wang H, et al. A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. ACM Trans Inf Syst. 2025;43(2):1-55. doi: 10.1145/3703155
5. Paladines J, Ramirez J. A Systematic Literature Review of Intelligent Tutoring Systems With Dialogue in Natural Language. IEEE Access. 2020;8:164246-164267. doi: 10.1109/access.2020.3021383
6. de la Torre A, Baldeon-Calisto M. Generative Artificial Intelligence in Latin American Higher Education: A Systematic Literature Review. 2024 12th International Symposium on Digital Forensics and Security (ISDFS). 2024:1-7. doi: 10.1109/isdfs60797.2024.10527283
7. Betti Galasso BJ. Inteligencia artificial generativa y formación docente: Competencias emergentes en América Latina – una revisión de alcance. Act Inv en Educ. 2025;26(1):1-43. doi: 10.15517/z4mcs052
8. Torres-Diaz JC, Duart J, Rivera D, Flandoli AB. Artificial intelligence and academic integrity: exploring plagiarism in Ecuadorian universities. Int J Educ Integr. 2025;21(1):35. doi: 10.1007/s40979-025-00209-3
9. Ancheta-Arrabal A, Pulido-Montes C, Carvajal-Mardones V. Gender Digital Divide and Education in Latin America: A Literature Review. Education Sciences. 2021;11(12):804. doi: 10.3390/educsci11120804
10. Sailer M, Homner L. The Gamification of Learning: a Meta-analysis. Educ Psychol Rev. 2020;32(1):77-112. doi: 10.1007/s10648-019-09498-w
11. Zeng J, Sun D, Looi C, Fan ACW. Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from the year 2008 to 2023. Brit J Educational Tech. 2024;55(6):2478-2502. doi: 10.1111/bjet.13471
12. Gatti Junior W, Marasco E, Kim B, Behjat L, Eggermont M. How ChatGPT can inspire and improve serious board game design. IJSG. 2023;10(4):33-54. doi: 10.17083/ijsg.v10i4.645
13. Montella R, De Vita CG, Mellone G, Ciricillo T, Caramiello D, Di Luccio D, et al. Leveraging Large Language Models to Support Authoring Gamified Programming Exercises. Applied Sciences. 2024;14(18):8344. doi: 10.3390/app14188344
14. Alharthi SA. Generative AI in Game Design: Enhancing Creativity or Constraining Innovation?. J Intell. 2025;13(6):60. doi: 10.3390/jintelligence13060060
15. Al-Surayhi MM, Al-Qahtani MR, Al-Marshadi SN, Al-Ghamdi AAQ, Al-Surayhi MM. Generative Artificial Intelligence in Game Design: A Narrative Review. JoE. 2025;4(4). doi: 10.62754/joe.v4i4.6912
16. de la Puente Pacheco MA, Torres J, Blanco Troncoso AL, Guzmán Murillo HJ, Carrascal JXM. Enhancing Critical Thinking and Argumentation Skills in Colombian Undergraduate Diplomacy Students: ChatGPT-Assisted and Traditional Debate Methods. Journal of Political Science Education. 2025;21(4):728-738. doi: 10.1080/15512169.2025.2449936
17. Avello-Martínez R, Gajderowicz T, Gómez-Rodríguez VG. Is ChatGPT helpful for graduate students in acquiring knowledge about digital storytelling and reducing their cognitive load? An experiment. Red. 2024;24(78). doi: 10.6018/red.604621
18. Mellado R, Cubillos C. Can Generative Artificial Intelligence Outperform Self-Instructional Learning in Computer Programming?: Impact on Motivation and Knowledge Acquisition. Applied Sciences. 2025;15(11):5867. doi: 10.3390/app15115867
19. Gomes LN, Fraga CF, Holanda M, Da Silva D. ChatGPT as a Teaching Assistant in a CS1 course: An Experience Report. 2025 IEEE Frontiers in Education Conference (FIE). 2025:1-9. doi: 10.1109/fie63693.2025.11328432
20. Rojas Guillen JA, Gamarra Moreno AH, Quispe Bautista WE, Palacios Rojas DM. Impact of the ROCEF Methodology on the development of critical thinking in university students through Artificial Intelligence. Proceedings of the 23rd LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI): "Engineering, Artificial Intelligence, and Sustainable Technologies in service of society". 2025;(2025). doi: 10.18687/laccei2025.1.1.506
21. Gonzales JDR, Reyna JT, Amaya EAV, Luján IGV. Evaluation of the Impact of ChatGPT on the Development of Research Skills in Higher Education. IJLTER. 2025;24(4):370-390. doi: 10.26803/ijlter.24.4.18
22. Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Monterrey, México, Antonio CH. Examining the Impact of ChatGPT on Mexican Students' Performance in Higher Education. IJIET. 2025;15(7):1368-1376. doi: 10.18178/ijiet.2025.15.7.2338
23. Melgarejo-Solis R, Vargas-Chota P, Núñez-Alcántara A, Oviedo-Vargas S, Pujay-Huanca M, Romero-Unteveros L, et al. Artificial Intelligence (Chat GPT) and Its Impact on the Academic Success of Software Engineering Students at the UNMSM. Proceedings of the 5th LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2025): "Entrepreneurship with Purpose: Social and Technological Innovation in the Age of AI". 2025. doi: 10.18687/leird2025.1.1.557
24. Carabeo RJ, Castro LA. Designing a Generative AI–Powered Virtual Tutor: Studying Interaction Modality and Avatar Presence on the User Experience. Avances en IHC. 2025;10(1):52-59. doi: 10.47756/aihc.y10i1.201

25. Iglesias RFC, Gonzales AA, Durango DWB, Saenz CAT. Architecture and Evaluation of a Generative AI Chatbot for Personalized Learning in Secondary Education. 2025 IEEE 5th International Conference on Advanced Learning Technologies on Education & Research (ICALTER). 2025:1-4. doi: 10.1109/icalter69698.2025.11355110
26. Mejia YZOA, Navarro CAD, Casildo-Bedón NE, Pezo YER. Influence of a chatbot based on a conversational agent on the adaptability of first-year students of a Peruvian private university. Front Educ. 2024;9:1459303. doi: 10.3389/feduc.2024.1459303
27. Soriano KT, Inga-Avila MF, Churampi-Cangalaya RL. Impact of intelligent tutoring on emotion and academic performance of systems engineering students at the national university of central Peru. 105267/jijdns. 2025;9(2):289-296. doi: 10.5267/j.ijdns.2024.11.002
28. Guzman Ramos GG, Calderon M. Teaching Computational Thinking: A Tangible Card–Based Approach with Multimodal Large Language Models and TiniScript. 2025 44th International Conference of the Chilean Computer Science Society (SCCC). 2025:1-7. doi: 10.1109/sccc67219.2025.11420641
29. Ossa C, Willatt C. Uso de Inteligencia Artificial Generativa para retroalimentar escritura académica en procesos de Formación Inicial Docente. Eur j educ psicol. 2023;16(2):1-16. doi: 10.32457/ejep.v16i2.2412
30. Bustillos A, Marizande F, Cevallos A, Bustillos D, Arteaga C, Vásquez de la Bandera F. Evaluation of the efficacy of ChatGPT versus medical students in clinical case resolution. Data and Metadata. 2024;3:433. doi: 10.56294/dm2024.433
31. Carrasco-Sáez JL, Contreras-Saavedra C, San-Martín-Quiroga S, Contreras-Saavedra CE, Viveros-Muñoz R. Analyzing Higher Education Students' Prompting Techniques and Their Impact on ChatGPT's Performance: An Exploratory Study in Spanish. Applied Sciences. 2025;15(14):7651. doi: 10.3390/app15147651
32. Mora Preciado FA, Molina Pinargote FM. Impact of ChatGPT's Conversational Features on EFL Learners' English-Speaking Proficiency in Latin America: A Literature Review. Arandu. 2025;12(3):900-913. doi: 10.69639/arandu.v12i3.1353
33. Guerrero-Zambrano M, Sanchez-Alvarado L, Valarezo-Chamba B, Lamilla-Rubio E. Transforming Physics Teacher Training Through ChatGPT: A Study on Usability and Impact. Education Sciences. 2025;15(7):887. doi: 10.3390/educsci15070887
34. Cordero Delgado LE, Cabal Hidalgo FDR, Nevárez Mieles EA, Matamoros Dávalos ÁA. Impacto de la integración de la inteligencia artificial y la gamificación en el proceso de enseñanza-aprendizaje en la formación de docentes de educación inicial. Rev Cient Multi G-nerando. 2025;6(2). doi: 10.60100/rcmg.v6i2.873
35. Estrella Rocafuerte AM, Abad Troya L, Nancy Maribel MC, Alvarez Leon DS, Naula Pullo JC, Rivadeneira Ormaza JM. Aplicación de la Inteligencia Artificial y gamificación para el Aprendizaje para Estudiantes con Necesidades Educativas Específicas en lenguaje. Vitalia. 2025;6(2):382-419. doi: 10.61368/r.s.d.h.v6i2.608
36. Aguirre Pérez GR, Aguirre Pérez MDLM, Ortega Acosta Juan Carlos Enrique OAJCE, Izquierdo Morán AM, Franco Castro Sara Alexandra SA. Influencia de gamificación e inteligencia artificial generativa en el proceso de enseñanza-aprendizaje en la Educación Básica Superior. COCIRI. 2025;6(2):1731-1759. doi: 10.55813/gaea/ccri/v6/n2/1271
37. Pulgarín Izquierdo YM, Espinoza Galván CV, Robles Cando ME, Castro Torres JR, Castro Quizhpi MV. Inteligencia artificial y gamificación como recursos inclusivos para la Educación Fiscal en el Ecuador. Epistemología de las Ciencias. 2025;2(3):358-377. doi: 10.71112/g28s3x45
38. Quinga M, Morán L, Fernández L, Vergel E. Gamificación con inteligencia artificial para mejorar la enseñanza de ciencias naturales en educación básica elemental. REG. 2025;4(4). doi: 10.70577/reg.v4i4.432
39. Avelar EA, Durso SDO, Ferreira CDO. Gamificação aplicada à educação contábil com base em algoritmos de inteligência artificial. REPeC. 2025;19. doi: 10.17524/repec.v19.e3615
40. Gomes Pontes A, Gomes Xavier MA, Lopes Xavier B, Rodrigues de Carvalho PV. Uma análise sobre aprendizagem adaptativa e Inteligência Artificial em jogos de quiz: Rapid Review. H2D. 2025;7(1):e6936. doi: 10.21814/h2d.6936
41. Guamán Toca FC, Iturralde Guerra LY, León Sampedro FG, Loor Avila BA. Gamificación en Educación Superior: Barreras de Adopción Docente y Estrategias para su Integración Efectiva. Veritas. 2025;6(3):877-896. doi: 10.61616/rvdc.v6i3.974
42. Camargo Arboleda MA, Herrera Rocha J. El impacto de la Inteligencia Artificial y la Gamificación en la Educación Contable. ASCE. 2025;4(1):120-133. doi: 10.70577/asce/120.133/2025
43. Casa Tinoco DL, Montero Murrieta EC, Avello-Martínez R, Tapia-Bastidas T. Formación docente en diseño de materiales interactivos con inteligencia artificial. minerva. 2025;6(11):227-256. doi: 10.53591/6nvhs134
44. Mora Villamar FM, Dume Naranjo IG, Urrutia Franco NY, Pozo Vera EI. Innovación Didáctica e Inclusión Digital en Lengua y Literatura: Aplicaciones de Tecnologías Emergentes, Gamificación e Inteligencia Artificial en el Aula del Siglo XXI. ASCE. 2025;4(4):1349-1376. doi: 10.70577/asce.v4i4.493
45. Canedo E, Calazans A, Silva G, Masson E, Mendonça F. Teaching Practice Using ChatGPT in Higher Education. Proceedings of the 26th International Conference on Enterprise Information Systems. 2024;1:541-552. doi: 10.5220/0012725500003690
46. Sánchez Molano B, López García JC. Edutekalab: inteligencia artificial al servicio del desarrollo de competencias del siglo XXI y el apoyo docente. Voces silec rev latinoam educ. 2025;16(2):192-214. doi: 10.18175/vys16.2.2025.10
47. Vargas Bernuy JB, Nolasco-Mamani MA, Velásquez Rodríguez NC, Gambetta Quelopana RL, Martinez Valdivia AN, Espinoza Vidaurre SM. Relative Advantage and Compatibility as Drivers of ChatGPT Adoption in Latin American Higher Education: A PLS SEM Study Towards Sustainable Digital Education. Sustainability. 2025;17(18):8329. doi: 10.3390/su17188329

48. Ballesteros MA, Acosta-Enriquez BG, Valle MDLÁG, Morales-Angaspilco JE, Callejas Torres JC, Luján López JE, et al. The influence of social norms and word-of-mouth marketing on behavioral intention and behavioral use of generative AI chatbots among university students. Computers in Human Behavior Reports. 2025;19:100760. doi: 10.1016/j.chbr.2025.100760
49. Cervantes D, Morales E. Perception and use of ChatGPT among university students. Analysis based on K-Modes. Data and Metadata. 2025;4:729. doi: 10.56294/dm2025729
50. Aguilar Ponce YA. Uso de chatbots de inteligencia artificial generativa en actividades de aprendizaje en estudiantes universitarios. Rev Cie Ava Cie Doc. 2025;2(Especial):35-43. doi: 10.70939/revistadiged.v2iespecial.49
51. Gómez Gómez J, Salas Álvarez D, Hernández Riaño V. A study on the perception of Generative AI in higher education students and teachers. inycomp. 2025;27(3). doi: 10.25100/iyc.v27i3.15137
52. Gonzales JDR, Reyna JT, Pajuelo JIT, Gomez RJM, Zavaleta ARS. Analyzing TPACK and Demographic Influences on Faculty Adoption of ChatGPT in Latin American Higher Education. IJLTER. 2025;24(7):402-424. doi: 10.26803/ijlter.24.7.20
53. Palacios-Núñez ML, Mendoza-García EM, Narciso Zarate JW, Deroncele-Acosta A. ChatGPT in the Teaching of Academic Writing in Higher Education: Teachers' Perspectives on Its Uses, Challenges, and Future in Personalized Learning. Edutec. 2025;2025(93):33-50. doi: 10.21556/edutec.2025.93.3995
54. Ortiz Garcia LK. Enhancing University Lecturers' Digital Competence through Generative Artificial Intelligence: A Quasi-Experimental Study in Peru. Proceedings of the 5th LACCEI International Multiconference on Entrepreneurship, Innovation and Regional Development (LEIRD 2025): "Entrepreneurship with Purpose: Social and Technological Innovation in the Age of AI". 2025. doi: 10.18687/leird2025.1.1.529
55. Castillo-Cuesta L, Ochoa-Cueva C, Cabrera-Solano P. Pre-service Teachers' Perceptions of the Use of Artificial Intelligence in an English as a Foreign Language Learning Context. IJLTER. 2025;24(10):818-841. doi: 10.26803/ijlter.24.10.39
56. Cardoso García D, Mero Salazar L, Saltos Arias D, Sánchez Sánchez M, Espinoza Mogrovejo M. La inteligencia artificial generativa y su aporte en la enseñanza personalizada para estudiantes de educación secundaria. LATAM. 2025;6(4). doi: 10.56712/latam.v6i4.4564
57. Calero-Mieles VM, Guerrero-Bermúdez ÁE, Barban-Forte Y. Inteligencia artificial generativa y aprendizaje crítico en educación secundaria. RevInvLid. 2024;1(1):56-70. doi: 10.63618/omd/revinvlid/v1/n1/10
58. Polastri PCP, Linhalis F, Dos Reis JC. Compreendendo o uso da inteligência artificial generativa por professores da educação básica no Brasil. CuadActiva. 2025. doi: 10.53995/20278101.2045
59. Vásquez Cardona F, Carmona García V, Pérez Valencia D. ChatGPT in Higher Education: A Cross-Sectional Study of Student Usage, Ethical Perceptions, and Implications for Academic Integrity. RBD. 2025;(65):101-119. doi: 10.1344/rbd2025.65.49416
60. Nelson AS, Santamaría PV, Javens JS, Ricaurte M. Students' Perceptions of Generative Artificial Intelligence (GenAI) Use in Academic Writing in English as a Foreign Language. Education Sciences. 2025;15(5):611. doi: 10.3390/educsci15050611
61. Amaya EAV, Zavaleta ARS, Luján IGV, López EG, Reyna JT, Gonzales JDR. Between Assistance and Dependence: Artificial Intelligence and Critical Thinking. A Study on First-Year University Students. IJLTER. 2025;24(12):118-134. doi: 10.26803/ijlter.24.12.6
62. Crisostomo RA, Sanchez C, Nunez AM, Nunez NA. More AI, Less Learning? The Role of Technological Self-Efficacy in STEM Education. 2025 International Symposium on Accreditation of Engineering and Computing Education (ICACIT). 2025:1-6. doi: 10.1109/icacit66767.2025.11369511
63. Buele J, Sabando-García ÁR, Sabando-García BJ, Yánez-Rueda H. Ethical Use of Generative Artificial Intelligence Among Ecuadorian University Students. Sustainability. 2025;17(10):4435. doi: 10.3390/su17104435
64. Levy Yeyati E, Robano V, Pereiro E, Porto C, Koleszar V. Generative AI in Education: A Framework for Leveraging Digital Tools in Latin American Classrooms [Internet]. Inter-American Development Bank; 2025. Available from:
65. Agüero Del Carpio RA, Agüero Del Carpio WH, Agüero Del Carpio PI, Agüero Del Carpio LE, Agüero Del Carpio RM, Bravo Huaynates GJ, et al. Technology and Education in the Era of Generative AI: Ethical-Social Analysis from University Classrooms in Peru. Proceedings of the 23rd LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI): "Engineering, Artificial Intelligence, and Sustainable Technologies in service of society". 2025;(2025). doi: 10.18687/laccei2025.1.1.1744
66. Jussara Reis-Andersson. AI Generative in Brazil's Public Schools: The Teachers' Perspective. ICAIR. 2024;4(1):374-380. doi: 10.34190/icair.4.1.3219
67. Rivadeneira L, Bellido De luna D, Fernandez C. Exploring the role of ChatGPT in higher education institutions: Where does Latin America stand?. Digit Gov: Res Pract. 2025;6(2):1-12. doi: 10.1145/3689370
68. Roberto Bernardo Usca Veloz, Wilian Alberto Yánez Arteaga, Edwin Olmedo Chávez Gavilánez, José Luis Vásconez Salazar. Generative Artificial Intelligence Systems and the Challenges in Latin American Education. esiculture. 2024:1210-1229. doi: 10.70082/esiculture.vi.1352
69. Kestin G, Miller K, Klales A, Milbourne T, Ponti G. AI tutoring outperforms in-class active learning: an RCT introducing a novel research-based design in an authentic educational setting. Sci Rep. 2025;15(1):17458. doi: 10.1038/s41598-025-97652-6
70. Atianashie MA, Kuffour MK, Kyiewu B, Clinton A. The impact of artificial intelligence based learning tools on students performance and engagement in higher education. SAP Social AI. 2025;1:44. doi: 10.62486/sai202544
71. Létourneau A, Deslandes Martineau M, Charland P, Karran JA, Boasen J, Léger PM. A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. npj Sci Learn. 2025;10(1):29. doi: 10.1038/s41539-025-00320-7

72. Urzúa CAC, Ranjan R, Saavedra EEM, Badilla-Quintana MG, Lepe-Martínez N, Philominraj A. Effects of AI-Assisted Feedback via Generative Chat on Academic Writing in Higher Education Students: A Systematic Review of the Literature. Education Sciences. 2025;15(10):1396. doi: 10.3390/educsci15101396
73. Fan Y, Tang L, Le H, Shen K, Tan S, Zhao Y, et al. Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. Brit J Educational Tech. 2025;56(2):489-530. doi: 10.1111/bjet.13544
74. Fang L, Tang G, Zhang L. Helpful or Harmful? Comparative Study of Perceived and Actual Effectiveness of LLM-Driven Tutors in Game-Based CFL Learning. Education Sciences. 2025;15(11):1502. doi: 10.3390/educsci15111502
75. Liang W, Yuksekgonul M, Mao Y, Wu E, Zou J. GPT detectors are biased against non-native English writers. Patterns. 2023;4(7):100779. doi: 10.1016/j.patter.2023.100779
76. García-López IM, Trujillo-Liñán L. Ethical and regulatory challenges of Generative AI in education: a systematic review. Front Educ. 2025;10:1565938. doi: 10.3389/feduc.2025.1565938

FUNDING

No financing

CONFLICT OF INTEREST

None

AUTHORSHIP CONTRIBUTIONS

Drafting – original draft: Analía Ortigoza, Lucía Gerez.
Writing–review and editing: Analía Ortigoza, Lucía Gerez.