

Impact of Chat GPT on Student Learning Outcomes: A Systematic Research Perspective on Academic Performance, Engagement, Critical Thinking, Self-Regulation, and Ethical Learning

Bathini Lavanya

Faculty in Computer Science & Applications, Govt. Degree College, Malkajgiri, Medchal Malkajgiri Dist., Telangana

<https://doi.org/10.55948/IJERMCA.2026.0902>

ABSTRACT

The rapid proliferation of generative artificial intelligence (Gen AI), particularly Chat GPT, has transformed higher education by reshaping how knowledge is produced, assessed, and mediated through digital systems. While these tools offer new opportunities for learning support and instructional efficiency, they also raise significant concerns regarding academic integrity, institutional governance, and ethical sustainability. The rapid diffusion of ChatGPT and other generative artificial intelligence (GenAI) tools has created a major transformation in educational practice. ChatGPT can provide explanations, examples, personalized feedback, language support, brainstorming, formative assessment and conversational tutoring at low marginal cost. At the same time, concerns have emerged about inaccurate information, over-reliance, reduced independent effort, academic integrity, unequal access, privacy and the possibility that students may obtain correct answers without developing the underlying knowledge and skills. This article examines the impact of ChatGPT on student learning outcomes, with particular attention to academic achievement, engagement, critical thinking, higher-order thinking, self-regulation, motivation, language development and metacognition.

The study adopts a systematic conceptual review methodology based on recent empirical studies, systematic reviews and meta-analyses published primarily between 2023 and 2026, supplemented by authoritative guidance from UNESCO. Recent experimental evidence is encouraging but heterogeneous. A 2025 meta-analysis of 69 experimental studies reported positive effects on academic performance, affective-motivational states and higher-order thinking, while finding no significant effect on self-efficacy. A separate 2025 meta-analysis of 57 studies and 97 estimates reported a large overall effect of GenAI on university learning outcomes, with especially strong effects for language skills, academic achievement, affective-motivational outcomes and higher-order thinking, but no significant effect on metacognition. A 2026 meta-analysis focused on undergraduate students reported a large positive overall effect for ChatGPT interventions, while emphasizing the need for contextual interpretation. These findings suggest that ChatGPT can enhance learning when integrated as a pedagogical support rather than used as an answer-substitution mechanism. The article identifies the conditions under which ChatGPT is more likely to improve learning and proposes institutional recommendations for AI literacy, assessment redesign, academic integrity, teacher training, equitable access and responsible use. The principal conclusion is that the educational impact of ChatGPT depends less on the technology alone than on the instructional design, type of task, student AI literacy, level of supervision and assessment structure.

Keywords: ChatGPT; Generative AI; Student Learning Outcomes; Academic Achievement; Critical Thinking; Higher-Order Thinking; Student Engagement; Self-Regulated Learning; Higher Education; AI Literacy; Academic Integrity.

INTRODUCTION

Generative artificial intelligence (Gen AI), particularly large language models (LLMs) such as Chat GPT, is rapidly transforming higher education by reshaping how knowledge is produced, assessed, and mediated through digital systems.

Launched by Open AI in November 2022, Chat GPT surpassed one million users within days, signaling strong student interest and institutional attention. Recent survey data illustrate this rapid uptake: 91% of university leaders believe GenAI will enhance and personalize learning, yet faculty adoption continues to lag behind student usage. Nearly half of leaders expect net institutional benefits, whereas a notable minority remain cautious about academic integrity and pedagogical disruption.

Despite this rapid diffusion, the role of Chat GPT in higher education remains contested. Reviews warn that Gen AI may undermine academic integrity, even as they acknowledge benefits for assessment and feedback. Empirical findings reflect similar tensions: some studies report improvements in critical and creative thinking, problem-solving, and motivation, while others document null or negative effects on learning performance and higher-order thinking. Lim et al. (2023) note that GenAI is both disruptive and potentially enriching, suggesting that educational technologies often have multiple and context-dependent influences.

The International Center for Academic Integrity (2024) reported that 58% of students admitted to using AI tools to complete assignments dishonestly, highlighting the urgency of addressing ethical concerns. Recent systematic reviews confirm that generative AI exacerbates academic integrity challenges by complicating authorship and enabling surface-level performance that may obscure limited cognitive engagement. Furthermore, fragmented policy responses—ranging from prohibition to cautious integration—create confusion among students about permissible and impermissible uses.

The Chat GPT marked a significant turning point in the relationship between artificial intelligence and education. Unlike earlier educational software that generally performed narrowly defined tasks, large language models can generate explanations, examples, summaries, questions, feedback and conversational responses across a wide range of subjects. As a result, students can interact with an AI system in a manner resembling an always-available study assistant. The educational possibilities are substantial, but so are the pedagogical and ethical questions.

Higher education has been one of the earliest environments in which ChatGPT has been adopted extensively. A 2024 systematic literature review identified 57 studies on ChatGPT in higher education and found applications involving students, academic staff, researchers and non-academic users. [1] A 2025 systematic review of 39 studies found that ChatGPT was being used for tutoring, essay drafting, content clarification and feedback, while also identifying concerns about academic integrity, critical thinking, over-reliance and technostress. [2]

The central issue, however, is not simply whether students use ChatGPT. The more important research question is whether and under what conditions ChatGPT changes what students actually learn. A student may produce a higher-quality essay with AI assistance without acquiring equivalent writing or reasoning skills. Conversely, a student who uses ChatGPT to obtain alternative explanations, generate practice questions, receive feedback and challenge their own reasoning may develop stronger learning outcomes. Therefore, educational research needs to distinguish between performance with AI assistance and durable learning that remains after the assistance is removed.

Recent experimental syntheses provide increasingly useful evidence. Deng et al. reviewed 69 experimental studies and reported that ChatGPT interventions improved academic performance, affective-motivational states and higher-order thinking propensities, while reducing mental effort; they found no significant effect on self-efficacy. [3] Chen and Cheung's 2025 meta-analysis synthesized 57 studies and 97 estimates and reported a large overall effect of GenAI on university learning outcomes, including language skills, academic achievement, affective-motivational status and higher-order thinking, but no statistically significant effect on metacognition. [4] A 2026 meta-analysis of undergraduate ChatGPT interventions reported a large positive overall effect based on 66 studies and 129 effect sizes. [5] These findings are promising, but they do not imply that unrestricted ChatGPT use automatically improves education. Effects vary by subject, intervention duration, learner characteristics, assessment type and instructional context. Recent research also emphasizes the importance of distinguishing AI-supported performance from independent learning and of using more authentic assessments that require students to demonstrate skills. [3,4]

Significance of the Study

The significance of the study lies in bringing together evidence on academic performance and broader learning outcomes. Student learning is multidimensional: examination scores, conceptual understanding, critical thinking, problem-solving, motivation, engagement, language proficiency, self-regulation and metacognition are not identical constructs. An educational technology can improve one dimension while having little or even negative influence on another.

Scope of the Study

The article focuses primarily on students in higher and post-secondary education, while drawing on broader GenAI evidence when it helps explain mechanisms. The review covers academic achievement, engagement, motivation, higher-

order thinking, critical thinking, language learning, self-regulation, metacognition, academic integrity and equity. It does not treat AI-generated text as inherently educational; rather, it examines how instructional design mediates its effects.

CONCEPTUAL FOUNDATIONS

ChatGPT as an Educational Technology

ChatGPT is a generative language model that produces responses based on patterns learned from large-scale training data and the context of a conversation. In education, its distinctive feature is interactivity. Students can ask follow-up questions, request examples, challenge explanations, ask for simplified language, generate practice exercises and obtain immediate feedback. This creates opportunities for individualized learning that traditional classroom instruction may not always provide at scale.

Student Learning Outcomes

Learning outcomes can be grouped into cognitive, affective, behavioral and metacognitive dimensions. Cognitive outcomes include knowledge acquisition, comprehension, application, analysis, evaluation and creation. Affective outcomes include motivation, confidence and attitudes toward learning. Behavioral outcomes include engagement, persistence and participation. Metacognitive outcomes include planning, monitoring, evaluation and awareness of one's own learning. Recent meta-analytic evidence indicates that GenAI effects differ across these dimensions, making multidimensional evaluation essential. [4,6]

Constructivist Learning Perspective

From a constructivist perspective, learning is not merely the receipt of information; students construct knowledge through interaction, prior experience, questioning, problem-solving and reflection. ChatGPT can support this process when it functions as a conversational scaffold. For example, students can ask the system to provide a counterargument, identify weaknesses in a draft, or generate progressively difficult practice problems. The same tool can undermine constructivist learning if students simply copy generated answers without processing them.

Self-Regulated Learning

Self-regulated learning involves goal setting, strategic action, monitoring and reflection. ChatGPT can support each stage: students can ask for study plans, practice tests, explanations and feedback. However, if students outsource planning and evaluation completely, the tool may reduce opportunities to practice self-regulation. This distinction is consistent with evidence that GenAI can improve many learning outcomes while effects on metacognition are weaker or nonsignificant. [4,6]

REVIEW OF LITERATURE

Study / source	Evidence base	Main findings	Implication
Baig & Yadegaridehkordi (2024) [1]	57 studies	ChatGPT applications are diverse across higher education.	Research is expanding rapidly; pedagogical context matters.
Systematic review (2025) [2]	39 studies, 2023–2024	Benefits for accessibility, engagement and support; concerns about integrity and over-reliance.	Balanced institutional guidance is needed.
Deng et al. (2025) [3]	69 experimental studies	Positive effects on academic performance, motivation and higher-order thinking; no significant self-efficacy effect.	Experimental evidence is positive but methodological limitations remain.
Chen & Cheung (2025) [4]	57 studies, 97 estimates	Large GenAI effect on university outcomes; language, achievement and HOTS benefits; no significant metacognition effect.	Outcomes differ by domain and context.
Jin & Sercu (2025) [7]	21 experimental studies	ChatGPT generally effective for knowledge	Medium-term interventions appear more promising

		acquisition; skills-development effects vary.	than short-term use.
Qian (2025) [8]	Systematic review	GenAI applications increasingly embedded in teaching and learning.	Pedagogical applications are broadening.
Daniel (2025) [9]	158 studies	Most reviewed studies reported cognitive-skill improvements.	AI may support academic skill development when pedagogically structured.
Undergraduate meta-analysis (2026) [5]	66 studies, 129 effect sizes	Large positive overall effect of ChatGPT applications on undergraduate learning.	Strong evidence of potential, but generalization requires context.
Han et al. (2025) [6]	68 studies, 337 effect sizes	Moderate positive overall GenAI effect with high heterogeneity.	Effects depend strongly on educational context.

The literature shows a consistent but qualified pattern: ChatGPT and GenAI can improve learning outcomes, but the magnitude and durability of the effect depend on how the technology is integrated. This is especially important because many early studies relied on self-report surveys, short interventions or assessments closely aligned with the intervention. More rigorous designs require delayed post-tests, authentic performance tasks, proctored assessments and measures of independent performance. [3,4,7]

STATEMENT OF THE PROBLEM

While extensive research has demonstrated ChatGPT's potential for enhancing learning, a critical gap exists between academic prototypes and production-ready systems. Existing meta-analyses focus narrowly on specific outcomes (e.g., academic achievement, engagement) or educational levels (K-12 vs. higher education), lacking a holistic view of ChatGPT's impact across learning performance, perception, and higher-order thinking. Furthermore, few studies address the integration of emerging pedagogies (e.g., problem-based learning, collaborative learning) or technological moderators (e.g., model version, platform) that shape ChatGPT's efficacy. There is an urgent need for a comprehensive meta-analysis that bridges these gaps, providing researchers and practitioners with evidence-based guidance on when and how ChatGPT can be most effectively used in higher education practice

Objectives of the Study

1. To examine the impact of ChatGPT on students' academic learning outcomes.
2. To analyze its influence on academic achievement, engagement, motivation and higher-order thinking.
3. To examine whether ChatGPT supports critical thinking, problem-solving and language development.
4. To assess the potential effects of ChatGPT on self-regulated learning and metacognition.
5. To identify the pedagogical conditions under which ChatGPT is most likely to improve learning.
6. To examine risks involving over-reliance, misinformation, academic integrity and inequitable access.
7. To propose institutional and classroom strategies for responsible and effective ChatGPT integration.

RESEARCH METHODOLOGY

This article adopts a systematic conceptual review methodology. Recent empirical studies, systematic reviews and meta-analyses were identified through targeted searches of scholarly literature, with emphasis on studies that examined measurable learning outcomes rather than only attitudes toward AI. Authoritative policy guidance from UNESCO was also incorporated to address human-centered educational governance. [10]

Research component	Approach
Study design	Systematic conceptual review of empirical and review literature.
Main period	2023–2026, with foundational sources included where necessary.
Population	Primarily higher-education and undergraduate students.
Main exposure	ChatGPT / GenAI-supported learning.
Outcomes	Achievement, engagement, motivation, HOTS, critical thinking, language, self-regulation and metacognition.

Risk dimensions	Academic integrity, misinformation, over-reliance, privacy, equity and assessment validity.
Synthesis	Narrative comparison of recent systematic reviews, meta-analyses and empirical studies.

The review does not pool the underlying studies itself because recent meta-analyses already provide quantitative synthesis. Instead, it triangulates the results of major reviews and identifies convergent findings, contradictions and methodological limitations. This approach is appropriate for a research article focused on interpreting the educational meaning of current evidence.

IMPACT OF CHATGPT ON ACADEMIC ACHIEVEMENT

Academic achievement is one of the most frequently studied outcomes. Recent experimental syntheses generally report positive effects. Deng et al. found that ChatGPT interventions improved academic performance across the experimental literature. [3] Chen and Cheung reported an effect size of $g^+ = 0.633$ for academic achievement within their GenAI meta-analysis, while the overall learning-outcome effect was larger. [4] The 2026 undergraduate meta-analysis also reported a large overall positive effect of ChatGPT applications. [5]

Several mechanisms may explain these effects. First, ChatGPT can provide immediate explanations when students encounter difficulties. Second, it can generate additional examples and practice questions. Third, it can provide formative feedback before a final submission. Fourth, it can support language comprehension for students who are studying in a second language. Fifth, it can reduce the time required to locate introductory explanations, allowing students to spend more time on application and analysis.

However, improved assignment scores should not automatically be interpreted as improved learning. If students submit AI-generated answers without understanding them, observed performance may rise while independent competence remains unchanged. Therefore, assessments should distinguish between AI-assisted performance and unaided performance. Delayed tests, oral examinations, demonstrations, project defenses and process portfolios can provide stronger evidence of durable learning.

Personalized Academic Support

One of ChatGPT's most important educational advantages is conversational personalization. Students can ask for an explanation at a particular level, request an analogy, ask for additional examples or identify the precise step where they are confused. This can reduce barriers for students who hesitate to ask questions in a classroom setting.

Immediate Feedback

Immediate feedback can support formative learning by allowing students to identify errors before misconceptions become established. However, students need to evaluate the feedback because generative AI can produce plausible but incorrect statements. AI literacy must therefore accompany AI access.

Learning Accessibility

ChatGPT can provide multilingual explanations, simplified language, alternative formats and conversational assistance. This may benefit language learners and students who require additional explanation outside classroom hours. Recent evidence shows particularly strong effects of GenAI on language-related outcomes. [4]

IMPACT ON STUDENT ENGAGEMENT AND MOTIVATION

Engagement involves attention, participation, persistence and willingness to invest effort. ChatGPT can increase engagement by making learning interactive. Instead of passively reading an explanation, students can engage in dialogue, ask follow-up questions and request examples. Survey research has also reported positive associations between ChatGPT use and student engagement. [11]

Motivation can increase when students experience quick progress and reduced frustration. A student who is stuck on a concept can immediately request another explanation. However, motivation can also become externally dependent if students expect the AI to perform difficult cognitive work. The pedagogical goal should therefore be to use ChatGPT to sustain productive effort rather than eliminate effort.

Potential positive effect	Potential negative effect
Interactive dialogue	Passive acceptance of AI answers
Immediate feedback	Dependence on external feedback
Reduced frustration	Reduced tolerance for productive struggle
Personalized explanations	Over-personalization without teacher oversight
24/7 access	Increased screen/technology dependence
Practice-question generation	Superficial practice if answers are immediately revealed

IMPACT ON CRITICAL THINKING AND HIGHER-ORDER THINKING

Critical thinking is a central concern in debates about ChatGPT. Critics argue that students may outsource reasoning, while proponents argue that AI can be used to generate counterarguments, expose assumptions and support evaluation. The evidence increasingly suggests that structured interventions can improve higher-order thinking. Deng et al. reported positive effects on higher-order thinking propensities, and Chen and Cheung reported a medium effect for higher-order thinking. [3,4]

The distinction is between using ChatGPT as an answer engine and using it as an intellectual challenger. A teacher can instruct students to generate an initial answer independently, ask ChatGPT for an alternative position, compare the two, identify errors and defend a final conclusion. Such a sequence requires evaluation rather than copying.

- Ask ChatGPT to provide competing explanations rather than a single answer.
- Require students to identify assumptions and weaknesses in AI-generated reasoning.
- Ask students to verify factual claims using textbooks, peer-reviewed sources or official documents.
- Require students to produce a final argument in their own words and explain why they accepted or rejected AI suggestions.
- Use oral defense or reflective commentary to assess the student's reasoning process.

Impact on Language Learning and Academic Writing

Language learning is one area in which ChatGPT can provide particularly direct support. Students can request vocabulary explanations, grammar correction, conversational practice, translations, examples and feedback on drafts. The 2025 GenAI meta-analysis reported a particularly large effect for language skills. [4]

For academic writing, ChatGPT can help students brainstorm, structure an argument, identify unclear sentences and suggest revisions. The risk is that students may surrender authorship and writing practice. A productive model is therefore process-based writing: students submit an initial draft, receive AI-assisted feedback, revise it, and provide a reflection explaining the changes.

In this approach, ChatGPT becomes a writing tutor rather than a ghostwriter. Assessment can focus on planning, evidence selection, argument development, revision and oral explanation, all of which are more difficult to outsource than the final text alone.

Impact on Self-Regulated Learning and Metacognition

Self-regulated learning is one of the most important areas where evidence is mixed. ChatGPT can help students formulate study plans, generate practice questions, identify weak areas and reflect on mistakes. Yet the same convenience may reduce opportunities for students to plan and monitor their own learning.

The evidence is particularly noteworthy because Chen and Cheung found no statistically significant effect of GenAI on metacognition. [4] This suggests that students may learn more or perform better without necessarily becoming better at monitoring how they learn. The implication is that metacognitive activities should be deliberately designed rather than assumed to emerge automatically from AI use.

Stage of self-regulation	Possible ChatGPT support	Required student responsibility
Planning	Generate study schedule and learning goals.	Select realistic goals and justify priorities.
Monitoring	Ask for practice questions and diagnostic feedback.	Judge whether the feedback is accurate.
Strategy selection	Compare alternative study methods.	Choose a method based on evidence.
Reflection	Generate reflection prompts.	Explain what was learned and what

		remains uncertain.
Evaluation	Check draft against criteria.	Make final independent judgment.

Impact of Chat GPT on Higher Education Student Learning Outcomes

ChatGPT has a moderately positive impact on higher education student learning outcomes, with meta-analyses reporting effect sizes around $g \approx 0.58$ – 0.67 , indicating meaningful gains in academic performance and related skills when used appropriately. However, benefits are not automatic: they depend on how the tool is integrated into teaching, the subject area, duration of use, and whether safeguards against overreliance are in place..

What the evidence shows about learning outcomes

Overall academic achievement: A 2025 meta-analysis of empirical studies found a moderately positive effect of ChatGPT on students' academic achievement ($g = 0.577$, $p < 0.001$). A 2026 meta-analysis similarly reported a positive effect ($g = 0.670$), with improvements in both cognitive (e.g., understanding, problem-solving) and non-cognitive outcomes (e.g., engagement, confidence).

Task performance and cognitive load: Experimental work shows that students using ChatGPT-based aids had higher task completion, lower cognitive load, **and** better critical thinking in some contexts, suggesting it can offload routine processing and free cognitive resources for deeper work.

Perceptions vs grades: Large surveys indicate students *feel* ChatGPT improves efficiency, access to knowledge, and chances of good grades, but some studies find self-reported benefits do not always translate into higher exam scores, highlighting a perception–performance gap.

Where ChatGPT helps most (moderators and mechanisms)

Research identifies several conditions under which impacts are stronger:

- **Subject area:** Effects are larger in social sciences than in some other disciplines, possibly due to text-heavy tasks and conceptual explanation needs.
- **Duration and design:** Interventions of 5–10 weeks show larger impacts than very short or very long exposures, and blending ChatGPT with traditional classrooms outperforms some flipped-classroom-only designs.
- **Type of knowledge and output:** ChatGPT is more effective for declarative knowledge (facts, concepts, explanations) than procedural knowledge, and text generation tasks show stronger achievement gains than code generation in some analyses.
- **Mechanisms of benefit:** Students use ChatGPT to summarize complex material, get instant explanations, generate alternative examples, and receive personalized, on-demand feedback, which can improve understanding, writing quality, and self-efficacy.

Risks and negative pathways

Potential downsides are well documented and can offset gains if unmanaged:

- ❖ **Overreliance and shallow learning:** Heavy dependence may reduce effortful inquiry, analysis, and synthesis, risking **surface-level understanding** and weaker critical engagement.
- ❖ **Critical thinking concerns:** While some studies report improved critical thinking with guided use, others warn that bypassing core cognitive steps can **impede deeper learning** over time.
- ❖ **Academic integrity and ethics:** Issues include **plagiarism**, **ethical ambiguity**, and **content accuracy** (hallucinations), requiring clear policies and AI literacy

Practical implications for Higher education

Given your focus on economics/digital marketing and academic writing, the literature suggests high-leverage, low-risk uses:

- ✓ **Use as a tutor/explainer:** Ask Chat GPT to **re-explain theories** (e.g., trade models, public expenditure frameworks) in simpler terms, provide **counterexamples**, or generate **practice questions** with solutions for revision.
- ✓ **Writing support, not substitution:** Use it to **outline**, **clarify structure**, and **get feedback** on drafts, then rewrite in your own voice and verify citations—this preserves learning while gaining efficiency.
- ✓ **Structured integration:** Embed ChatGPT in **time-bound activities** (e.g., 5–10 week modules) with explicit prompts that require **analysis**, **comparison**, and **justification**, not just summary, to avoid overreliance.
- ✓ **AI literacy and assessment design:** Teach students to **fact-check**, **cite AI use**, and design assessments that value **process artifacts** (drafts, reflections, data work) over final text alone.

RISKS AND NEGATIVE EFFECTS

Over-Reliance

Over-reliance occurs when students use ChatGPT for tasks they could reasonably perform themselves. Excessive dependence can reduce practice in retrieval, writing, problem-solving and independent reasoning. The educational solution is not necessarily prohibition; it is purposeful task allocation and explicit boundaries around when AI assistance is permitted.

Hallucinations and Incorrect Information

ChatGPT can produce fluent and convincing but incorrect information. Students who lack sufficient subject knowledge may be unable to recognize errors. AI literacy should therefore include verification, source checking and uncertainty awareness.

Academic Integrity

Generative AI complicates conventional definitions of authorship and plagiarism. Institutions increasingly need to distinguish permitted assistance from unauthorized substitution of student work. UNESCO's guidance emphasizes a human-centered approach, appropriate governance, capacity building and consideration of ethical risks. [10]

Privacy and Data Protection

Students may enter personal information, unpublished research, confidential institutional material or assessment content into AI systems. Educational institutions should provide clear rules on what information may be shared and should teach students to minimize sensitive data exposure.

Inequality and Digital Divide

Students differ in access to devices, connectivity, paid AI features and AI literacy. If ChatGPT becomes an implicit requirement without institutional support, existing inequalities may widen. Equitable access and transparent alternatives are therefore important.

Reduced Productive Struggle

Learning often requires effortful retrieval and problem-solving. If ChatGPT provides an answer too early, students may lose opportunities to struggle productively. Teachers should design prompts and activities that require students to attempt the task before consulting AI.

CHATGPT AND ASSESSMENT REDESIGN

The growth of generative AI requires a shift in assessment design. Conventional take-home essays and factual questions can often be completed with AI assistance, making them weaker indicators of independent learning. Recent meta-analytic work specifically recommends more complex project-based assessment, proctored assessment, originality measures and objective outcomes. [3]

Traditional assessment	AI-resilient alternative
Take-home essay only	Draft + revision history + oral defense.
Short factual assignment	Application to a new case or local problem.
Generic homework	Personalized problem with explanation of reasoning.
Literature summary	Critical synthesis with source verification.
AI-generated answer accepted	AI critique and verification assignment.
Final examination only	Combination of formative tasks, projects and supervised assessment.

AI-resilient assessment does not mean assessment that assumes AI is absent. Instead, it assesses capabilities that remain educationally meaningful in an AI-rich environment: problem framing, evidence evaluation, judgment, creativity, communication, ethical reasoning and the ability to explain decisions.

Role of Teachers in the ChatGPT Era

ChatGPT changes the teacher's role but does not make teachers unnecessary. Teachers remain responsible for curriculum design, disciplinary judgment, assessment, motivation, social learning and ethical guidance. Their role increasingly includes teaching students how to interrogate AI outputs rather than merely preventing AI use.

- Design learning activities in which AI supports rather than replaces cognitive work.

- Teach students to verify AI-generated claims.
- Model appropriate prompting and responsible AI use.
- Explain institutional rules for disclosure and attribution.
- Use formative assessment to observe learning processes.
- Monitor unequal participation and access.
- Develop students' AI literacy alongside disciplinary knowledge.
-

Proposed Pedagogical Framework

This article proposes a five-stage CHATGPT-Learning Integration Framework: Challenge, Human attempt, AI interaction, Testing and Personal reflection.

Stage	Purpose	Example
C – Challenge	Present an authentic learning problem.	Analyze a policy problem.
H – Human attempt	Require independent first response.	Student drafts initial argument.
A – AI interaction	Use ChatGPT as tutor or critic.	Ask for counterarguments and feedback.
T – Testing	Verify AI claims and apply knowledge.	Check sources and solve a new problem.
P – Personal reflection	Demonstrate learning without AI dependence.	Explain what changed and why.

The framework deliberately places human reasoning before AI assistance and independent verification after it. This reduces the risk that students merely copy AI output and increases the opportunity to develop metacognitive awareness.

Expected Learning Pathway

Human attempt → AI feedback/dialogue → verification → revision → independent demonstration → reflection.

Empirical Research Model for Future Study

A future empirical study can test the framework using a quasi-experimental or randomized controlled design. Students can be divided into a control group receiving conventional instruction and an intervention group using structured ChatGPT activities. Pre-tests and post-tests should be combined with delayed post-tests to distinguish immediate performance from retention.

Variable	Measurement examples
Independent variable	Structured ChatGPT-supported learning vs conventional learning.
Academic achievement	Standardized course test / validated achievement test.
Critical thinking	Validated critical-thinking assessment and rubric.
Engagement	Behavioral participation + validated engagement scale.
Motivation	Validated motivation scale.
Self-regulation	Self-regulated learning scale + learning logs.
Metacognition	Metacognitive awareness inventory + reflection task.
Independent learning	No-AI delayed assessment.
AI literacy	Verification, prompting and ethical-use task.
Over-reliance	Frequency and type of AI use; substitution index.

Statistical analysis can include ANCOVA controlling for baseline achievement, mixed-effects models for repeated measures, mediation analysis to examine whether engagement or self-regulation explains achievement changes, and moderator analysis for prior achievement, subject area, intervention duration and AI literacy. Researchers should report effect sizes and confidence intervals rather than relying only on statistical significance.

DISCUSSION

The current evidence supports a cautiously positive interpretation of ChatGPT's educational impact. Experimental meta-analyses increasingly report improvements in academic performance and higher-order thinking, and the 2026 undergraduate

meta-analysis suggests a large positive overall effect. [3–5] However, the evidence also demonstrates that learning outcomes are heterogeneous. GenAI effects vary by educational level, subject, duration, assessment type and context. [4,6] This heterogeneity explains why both enthusiastic and negative classroom experiences can be valid. A student using ChatGPT as an interactive tutor may experience substantial benefit, while a student using it to complete assignments with minimal engagement may experience little durable learning. The same tool can therefore function either as a scaffold or as a substitute for cognition depending on task design.

A particularly important distinction is between performance enhancement and learning enhancement. ChatGPT can improve the quality of a submitted product even when the student has not acquired equivalent underlying skills. This is why delayed and AI-free assessments are important. The strongest educational model is one in which AI assistance is followed by independent demonstration.

The evidence on metacognition is more cautious. The lack of a significant effect in recent meta-analysis suggests that simply giving students access to ChatGPT does not automatically make them better learners. Metacognitive skills need explicit instructional treatment through planning, monitoring, self-explanation and reflection.

The future of ChatGPT in education should therefore be framed around augmentation rather than replacement. AI can expand feedback capacity, personalize explanations and support practice, while teachers preserve responsibility for disciplinary knowledge, assessment and educational relationships.

RECOMMENDATIONS

- Develop institution-wide generative-AI policies that distinguish acceptable assistance from unauthorized substitution.
- Introduce AI literacy as a core student skill, including prompting, verification, source evaluation, privacy and ethical use.
- Redesign assessments toward authentic, process-oriented and AI-resilient tasks.
- Use delayed or supervised assessments to measure independent learning.
- Require students to disclose meaningful AI assistance where appropriate.
- Use ChatGPT primarily for tutoring, feedback, brainstorming and practice rather than direct answer substitution.
- Teach students to challenge and verify AI outputs rather than treating them as authoritative.
- Provide equitable access or non-AI alternatives so that technology adoption does not widen educational inequality.
- Train teachers in pedagogical uses, assessment redesign and AI-related academic integrity.
- Evaluate learning outcomes over longer periods to separate novelty effects from durable learning.
- Collect evidence on metacognition and self-regulated learning rather than focusing only on grades.
- Use multidisciplinary institutional committees to review AI policy as models and educational practices evolve.

FUTURE RESEARCH DIRECTIONS

- Longitudinal studies examining whether ChatGPT-supported learning produces durable knowledge retention.
- Randomized studies comparing AI-assisted performance with delayed AI-free assessments.
- Research on ChatGPT's effect on metacognition and self-regulated learning.
- Cross-disciplinary studies comparing humanities, social sciences, science, engineering, medicine and vocational education.
- Research involving diverse socioeconomic and linguistic populations.
- Studies of AI literacy as a moderator of learning outcomes.
- Development of assessment models that measure reasoning, judgment and originality in AI-rich environments.
- Research on multimodal ChatGPT and emerging AI tutors beyond text-only interaction.
- Comparative research across different generative AI systems and versions.
- Studies of teacher workload and how AI changes instructional practice.

CONCLUSION

ChatGPT is one of the most consequential educational technologies to emerge in recent years. Current research increasingly indicates that it can improve academic achievement, engagement, motivation, language skills and higher-order thinking when used within structured learning interventions. Recent meta-analyses provide stronger evidence than early perception studies, although they also reveal substantial heterogeneity and methodological limitations.

The educational impact of ChatGPT should therefore not be reduced to a simple question of whether AI is good or bad for students. The more useful question is how the technology is used, for which tasks, by which learners, under what instructional conditions and with what assessment design. ChatGPT is most educationally valuable when it stimulates questioning, explanation, feedback, comparison, verification and reflection. It is least valuable when it replaces the student's own reasoning and production.

The most important policy implication is that educational institutions should move from prohibition-versus-permission debates toward AI literacy and responsible integration. Students need to know how to use ChatGPT, when not to use it, how to verify its output and how to demonstrate learning independently. Teachers need support in redesigning assignments and assessments. Institutions need clear rules on privacy, disclosure, academic integrity and equitable access.

In conclusion, ChatGPT has significant potential to improve student learning outcomes, but its effects are mediated by pedagogy. The goal should not be to create students who can obtain answers from AI; it should be to create students who can use AI intelligently while retaining the ability to think, evaluate, create, communicate and learn independently.

REFERENCES

1. Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in higher education: A systematic literature review and research challenges. *International Journal of Educational Research*, 127, 102411. <https://doi.org/10.1016/j.ijer.2024.102411>.
2. Systematic review of ChatGPT in higher education: Navigating impact on learning, wellbeing, and collaboration. (2025). *Social Sciences & Humanities Open*, 12, 101866. <https://doi.org/10.1016/j.ssaho.2025.101866>.
3. Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>.
4. Chen, S., & Cheung, A. C. K. (2025). Effect of generative artificial intelligence on university students learning outcomes: A systematic review and meta-analysis. *Educational Review*, 100737. <https://doi.org/10.1016/j.edurev.2025.100737>.
5. Undergraduate students' learning outcomes with ChatGPT: A meta-analytic study. (2026). *Computers and Education: Artificial Intelligence*, 10, 100536. <https://doi.org/10.1016/j.caeai.2025.100536>.
6. Han, X., Peng, H., & Liu, M. (2025). The impact of GenAI on learning outcomes: A systematic review and meta-analysis of experimental studies. *Educational Review*, 100714. <https://doi.org/10.1016/j.edurev.2025.100714>.
7. Jin, Y., & Sercu, L. (2025). ChatGPT interventions in higher education: A systematic review of experimental studies. *Journal of Computer Assisted Learning*, 41(4), e70072. <https://doi.org/10.1111/jcal.70072>.
8. Qian, Y. (2025). Pedagogical applications of generative AI in higher education: A systematic review of the field. *TechTrends*, 69, 1105–1120. <https://doi.org/10.1007/s11528-025-01100-1>.
9. Daniel, K. (2025). Can generative AI revolutionise academic skills development in higher education? A systematic literature review. *European Journal of Education*, 60(1), e70036. <https://doi.org/10.1111/ejed.70036>.
10. Miao, F., & Holmes, W. (2023). *Guidance for Generative AI in Education and Research*. UNESCO. Updated information accessed 2026.
11. Examining the effect of ChatGPT usage on students' academic learning and achievement: A survey-based study in Ajman, UAE. (2024). *Computers & Education: Artificial Intelligence*, 7, 100316. <https://doi.org/10.1016/j.caeai.2024.100316>.
12. Role of activity-based learning and ChatGPT on students' performance in education. (2024). *Computers & Education: Artificial Intelligence*, 6, 100219. <https://doi.org/10.1016/j.caeai.2024.100219>.
13. Sandu, R., Gide, E., & Elkhodr, M. (2024). The role and impact of ChatGPT in educational practices: Insights from an Australian higher education case study. *Discover Education*, 3, 71.
14. Artificial intelligence powering education: ChatGPT's impact on students' academic performance through the lens of technology-to-performance chain theory. (2024). *Journal of Applied Research in Higher Education*, 17(5), 1661–1679. <https://doi.org/10.1108/JARHE-04-2024-0179>.
15. Kasneci, E., Sessler, K., Küchemann, S., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>.
16. Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>.
17. Tlili, A., Shehata, B., Adarkwah, M. A., et al. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. <https://doi.org/10.1186/s40561-023-00237-x>.

18. UNESCO. (2021). AI and Education: Guidance for Policy-makers. UNESCO.
19. UNESCO. (2023). Guidance for Generative AI in Education and Research. UNESCO.
20. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16, 39.
21. Kasneci, E., et al. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274.
22. OpenAI. (2024–2026). ChatGPT product and model documentation. OpenAI.