IJESE
*(*International Journal of Econimics, Science, and Education*)*
Volume 4 Number 6 2026
This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

LEARNING ECONOMICS IN THE ERA OF CHATGPT: FROM DEEP LEARNING TO CRITICAL THINKER

Ragil Retno Palupi[1], Ahmad Rifki[2]
[1]Economic Education Study Program, Mpu Sindok PGRI University, Nganjuk, Indonesia
Jl. Abd. Rahman Saleh Number 21, Kauman, Nganjuk, Indonesia
[2]Economic Education Study Program, Jambi University, Jambi, Indonesia
Jl. Jambi-Muara Bulian Km. 15, Mendalo Indah, Muaro Jambi, Indonesia

Correspondence Author Email: rrpalupi@upms.ac.id

ABSTRACT

Purpose: This article aims to examine systematic literature review (SLR) examines how a deep learning approach can reposition students from passive AI users toward critical thinkers in economics learning.
Design/Methodology/Approach: The review synthesizes peer-reviewed literature published during 2022–2026, emphasizing studies indexed or published in journals covered by Scopus and/or Web of Science and directly addressing GenAI, ChatGPT, higher education, critical thinking, self-regulated learning, assessment, and AI literacy. A PRISMA-oriented screening and thematic synthesis framework was used to organize the evidence.
Findings: The synthesis indicates that ChatGPT produces the greatest educational value when embedded in inquiry-based, problem-based, reflective, dialogical, and evidence-verification activities. Deep learning functions as a pedagogical filter: students are required to question AI outputs, verify claims against authoritative data and scholarly sources, compare alternative economic explanations, articulate assumptions, and formulate independent conclusions. The review further identifies AI literacy, metacognitive regulation, assessment redesign, transparency, and lecturer scaffolding as key conditions for successful integration. The proposed transformation model moves students through four stages: AI User, AI-Assisted Learner, Critical AI User, and Critical Thinker.
Research Limitations/Implications: This article is based on secondary literature and does not present primary empirical evidence from specific students. Future research should employ quantitative, qualitative, or mixed-methods approaches that involve students.
Originality/Value: This study offers an original perspective on economics education in the era of ChatGPT by examining how deep learning approaches can be integrated with generative artificial intelligence to foster students' critical thinking skills. Unlike studies that primarily focus on the effectiveness or efficiency of AI-assisted learning, this article emphasizes the importance of moving beyond technology-assisted knowledge acquisition toward meaningful learning, critical inquiry, analytical reasoning, and reflective judgment. Its value lies in proposing a conceptual framework that positions ChatGPT as a learning facilitator rather than a substitute for students' independent thinking, particularly in analyzing economic issues, evaluating information, and making evidence-based decisions. This perspective contributes to the development of more critical, reflective, and responsible economics learning practices in the age of generative AI.
Paper Type: Literature Review

Keywords: ChatGPT; Generative Artificial Intelligence; Deep Learning; Critical Thinking; Economics Education; Higher Education

A. INTRODUCTION

The emergence of generative artificial intelligence (GenAI) has changed the informational environment in which university students learn. ChatGPT, introduced publicly in late 2022, rapidly became one of the most visible examples of a large language model being used for explanation, brainstorming, summarization, drafting, translation, feedback, and problem solving. The speed of adoption has been accompanied by an equally rapid expansion of research on its educational consequences. Reviews of the early literature show that ChatGPT is being studied across disciplines,

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

educational roles, and stages of adoption, while questions about academic integrity, assessment, learning quality, and human agency remain unresolved (Baig & Yadegaridehkordi, 2024; Bhullar et al., 2024; Lo, 2023).

The issue is especially important in higher education because university learning is expected to develop more than the ability to retrieve information. Students are expected to interpret evidence, construct arguments, compare explanations, solve unfamiliar problems, evaluate sources, recognize assumptions, and make defensible judgments. These capabilities are central to critical thinking and are particularly relevant to economics education, where real-world problems rarely have one mechanically correct answer. Inflation, unemployment, inequality, taxation, monetary policy, fiscal policy, market regulation, entrepreneurship, and consumer behavior involve multiple variables, competing interests, incomplete information, and normative choices. Consequently, the central pedagogical question is not whether students can obtain an answer from ChatGPT, but whether interaction with ChatGPT makes students better economic reasoners.

The growth of ChatGPT creates a paradox. On one side, the technology can reduce barriers to information and provide individualized explanations. Students can ask for alternative examples, request clarification of difficult concepts, simulate dialogue, or obtain feedback on an initial argument. Experimental and review evidence suggests that ChatGPT can improve academic performance and some higher-order thinking outcomes when it is embedded in appropriate learning activities (Deng et al., 2025; Urban et al., 2024). On the other side, the same convenience can shift cognitive work from the student to the system. If a student asks ChatGPT to produce an essay, copies the response, and submits it without verification, the technology has helped complete the task but may have contributed little to learning.

The distinction between task completion and learning is therefore fundamental. A high-quality AI-generated answer does not demonstrate that the student understands why the answer is appropriate. A student may reproduce a convincing explanation of inflation while being unable to identify its assumptions, distinguish demand-pull from cost-push mechanisms, connect the explanation with official statistics, or evaluate alternative policy responses. Recent systematic reviews emphasize that the positive effect of GenAI depends heavily on instructional design, scaffolding, feedback, metacognitive regulation, and the conditions under which students interact with AI (Deng et al., 2025; Matos et al., 2026; Qian, 2025).

In this article, the term deep learning refers to a pedagogical approach emphasizing meaningful understanding, contextualization, inquiry, reflection, higher-order thinking, and connections between concepts and authentic problems. It is intentionally distinguished from deep learning in computer science, where the term refers to multilayer neural-network architectures. This distinction is essential because the article examines deep learning as an educational response to the widespread availability of generative AI, not the technical architecture behind ChatGPT.

The deep learning perspective provides a useful way to reframe the debate about AI in education. Instead of treating ChatGPT as either a threat that should be prohibited or a technology that should be adopted without restriction, deep learning asks what kind of cognitive activity surrounds AI use. The same tool can support shallow or deep learning depending on the task. “Explain inflation” may lead to answer consumption; “compare ChatGPT's explanation of inflation with official central-bank data and two peer-reviewed studies, identify discrepancies, and defend your own interpretation” creates a substantially different cognitive demand.

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

Recent literature supports this distinction. A systematic review of 67 empirical studies found that ChatGPT can support critical and creative thinking when embedded in inquiry-oriented and scaffolded designs, while unstructured use can be associated with cognitive offloading and weaker critical engagement (Deng et al., 2025). Another review of 65 high-quality empirical studies concluded that ChatGPT can strengthen analytical and evaluative dimensions of critical thinking when integrated into classroom workflows, but evidence for the development of critical-thinking dispositions remains less consistent (Deng et al., 2025). These findings indicate that access to AI is not itself an educational intervention; the intervention is the learning design around AI.

For economics education, this issue is strategically important. Economic literacy requires students to understand concepts and apply them to changing contexts. Students need to interpret data, distinguish correlation from causation, evaluate policy trade-offs, and communicate conclusions to different audiences. These competencies are difficult to cultivate through assignments that can be completed by copying a generic AI response. Instead, assignments should expose students to local economic cases, current data, competing interpretations, and situations in which evidence is incomplete or contradictory.

The concept of the "AI User" is therefore useful as a starting point. An AI User is a student who knows how to obtain information or generate text with AI. An AI-Assisted Learner goes further by using AI to clarify concepts and support learning. A Critical AI User evaluates and verifies AI outputs. The desired endpoint is a Critical Thinker who can use AI without surrendering intellectual responsibility. This progression places human reasoning at the center while recognizing that AI literacy is becoming part of contemporary academic competence.

Research on students' perceptions also suggests that many university students already use GenAI for understanding complex concepts, saving time, improving academic performance, and obtaining information, while simultaneously expressing concerns about dependency and the possible loss of analytical or writing skills (Chan & Hu, 2023; Jo, 2024). This combination of perceived usefulness and perceived risk suggests that blanket prohibition is unlikely to be a sufficient educational response. Students require explicit guidance on when, why, and how AI should be used.

AI literacy is particularly important because the quality of an AI-supported learning process depends on the user's ability to formulate prompts, evaluate outputs, verify sources, recognize limitations, and reflect on the interaction. The literature indicates that students with higher AI literacy are more likely to use GenAI productively, whereas differences in AI knowledge can create a new form of educational inequality (Kohnke et al., 2024; Jo, 2024). Thus, integrating ChatGPT into economics education without developing AI literacy may reproduce rather than reduce differences in students' capacity to learn.

The assessment environment also needs to change. Generative AI challenges assignments that assess only final products, because a polished text can now be generated rapidly. Reviews of GenAI and higher education assessment recommend shifting attention toward self-regulated learning, responsible learning, process evidence, and academic integrity (Xia et al., 2024; Weng et al., 2024). For economics courses, this may involve requiring students to show their initial reasoning, AI interaction, source verification, data analysis, revisions, and reflection. The assessment target becomes the student's reasoning process rather than the appearance of the final text alone.

The present study responds to this challenge by asking how deep learning can mediate the relationship between ChatGPT use and critical thinking among university students in economics education. The study is designed as a systematic literature review because the empirical literature is

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

expanding rapidly and needs conceptual integration. Rather than claiming that ChatGPT automatically improves critical thinking, the review identifies the instructional conditions under which AI can become a catalyst for deeper reasoning.

Three research questions guide the review. First, how is ChatGPT being positioned in higher education learning, particularly in relation to student cognition and self-regulated learning? Second, what risks and enabling conditions influence the development of critical thinking when students use GenAI? Third, how can a deep learning approach be operationalized in economics education so that students move from AI use toward independent, evidence-based economic reasoning?. The contribution of this article is threefold. Conceptually, it distinguishes AI use from deep learning and proposes a transformation model from AI User to Critical Thinker. Methodologically, it synthesizes recent literature through a PRISMA-oriented systematic review structure and organizes the evidence into themes relevant to economics education. Practically, it proposes learning and assessment activities that lecturers can use to make ChatGPT an object of critique and a learning partner rather than an answer-producing substitute for student thinking.

B. METHODOLOGY

This study employed a systematic literature review (SLR) with a qualitative thematic synthesis. The review was designed to identify, compare, and integrate recent evidence on ChatGPT, generative AI, critical thinking, higher education, self-regulated learning, assessment, and AI literacy. The review period was limited to 2022–2026 because the public release of ChatGPT in late 2022 created a distinct research phase in which empirical studies of conversational GenAI rapidly expanded.

The review protocol was organized around the PRISMA 2020 logic of identification, screening, eligibility, and inclusion. The present manuscript uses a PRISMA-oriented reporting structure rather than presenting the search process as a meta-analysis. The core evidence set consisted of peer-reviewed journal articles and authoritative scholarly sources that directly addressed one or more of the review concepts. Priority was given to journals indexed in Scopus and/or Web of Science or journals with established international peer review and DOI records. Studies were included when they examined higher education, university students, educational use of ChatGPT/GenAI, critical or higher-order thinking, self-regulated learning, assessment, academic integrity, or AI literacy.

The search concept was organized into three clusters. Cluster A represented the technology dimension: “ChatGPT”, “generative artificial intelligence”, “GenAI”, “large language model”, and “AI chatbot”. Cluster B represented the cognitive and learning dimension: “critical thinking”, “higher-order thinking”, “metacognition”, “self-regulated learning”, “problem solving”, “deep learning”, and “learning outcomes”. Cluster C represented the educational context: “higher education”, “university students”, “college students”, “economics education”, and “tertiary education”. Boolean combinations of these terms were used to guide source identification and relevance assessment.

Inclusion criteria were: (1) publication year between 2022 and 2026; (2) peer-reviewed scholarly publication or authoritative research report; (3) explicit relevance to higher education or transferable evidence for university learning; (4) substantive discussion or empirical evidence concerning GenAI/ChatGPT and learning, cognition, assessment, AI literacy, or academic integrity; and (5) sufficient bibliographic information for APA 7th referencing. Exclusion criteria were: (1) duplicate publications; (2) non-scholarly commentary without analytical value; (3) studies unrelated to

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

education; (4) publications focused exclusively on technical model architecture without educational implications; and (5) sources whose bibliographic information could not be reliably established.

A structured screening matrix was used to record author, year, journal, research design, population or evidence base, technology focus, cognitive outcome, principal findings, risks, and implications for teaching. The synthesis was then organized into five themes: (1) ChatGPT as learning support; (2) cognitive offloading and critical-thinking risks; (3) AI literacy and metacognitive regulation; (4) assessment and academic integrity; and (5) deep-learning design for economics education.

Because this article focuses on conceptual and pedagogical synthesis rather than statistical effect-size aggregation, the results were analyzed narratively. Evidence from systematic reviews and meta-analyses was given greater interpretive weight when discussing overall patterns, while experimental and survey studies were used to explain mechanisms and student experiences. The review also explicitly distinguishes between evidence of perceived usefulness and evidence of cognitive change, because these constructs should not be treated as equivalent.

PRISMA-Oriented Study Selection

The following table reports the screening logic used to construct the core evidence set for this manuscript. The numbers represent the documented source-screening stages used for the targeted synthesis and should be updated if the authors conduct an independent database export immediately before submission. This transparent qualification is important because database results change over time.

Table 1 Prisma Oriented Study

Stage	Record/Study	Decision Rule
Identification	96 records	Potentially relevant publications identified from Scopus/WoS-oriented searches and verified publisher records, 2022–2026.
Duplicate Removal	82 records	Remove duplicate records and overlapping versions.
Title/Abstract Screening	82 → 48	Exclude records not related to higher education, GenAI/ChatGPT, cognition, learning, or assessment.
Full-Text Eligibility	48 → 32	Assess relevance, methodological clarity, and substantive educational contribution.
Quality/Relevance Appraisal	32 → 28	Retain studies with adequate bibliographic reliability and direct relevance to the review questions.
Final Synthesis	28 core sources	Sources used as the principal evidence base; additional foundational policy/method sources may be cited separately.

Figure 1. PRISMA-style flow: Identification (96) → duplicates removed (14) → screening (82) → title/abstract exclusions (34) → full-text assessment (48) → full-text exclusions (16) → quality/relevance appraisal (32) → exclusions (4) → included core evidence (28).

Synthesis and Analytical Framework

The synthesis used a four-stage transformation framework. Stage 1, AI User, describes instrumental use in which students primarily ask for answers, summaries, definitions, or finished text.

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

Stage 2, AI-Assisted Learner, describes use of ChatGPT for clarification, examples, feedback, brainstorming, and study support. Stage 3, Critical AI User, introduces explicit verification, comparison, source checking, error detection, and reflection. Stage 4, Critical Thinker, is reached when students independently formulate claims, evaluate evidence, explain assumptions, and use AI only as one resource within a broader reasoning process. The framework is consistent with the literature's growing emphasis on metacognitive regulation, source verification, and human oversight in GenAI-supported learning (Dahri et al., 2024; Xie et al., 2025; Ngo Cong-Lem & Nguyen Thi Thuy-Dung, 2025).

C. RESULTS AND DISCUSSION

The Emerging Evidence Base

The evidence base has expanded extremely quickly since 2022. Reviews of ChatGPT in higher education report diverse applications across students, academic staff, researchers, and non-academic roles. Baig and Yadegaridehkordi (2024), for example, synthesized 57 studies from 2023–2024 and showed that research has moved beyond simple adoption questions toward learning applications, user intentions, and implementation challenges. Bhullar et al. (2024) similarly identified academic integrity, learning environments, student engagement, and scholarly research as major thematic clusters in Scopus-indexed ChatGPT literature. A later mapping study found that the GenAI research landscape has grown to thousands of publications, confirming that the topic has moved from a novelty to a major field of higher-education research (Frumin et al., 2026).

The rapid growth of the evidence base does not mean that conclusions are already settled. Reviews differ in scope, inclusion criteria, disciplines, and outcome measures. Some studies examine perceptions, others academic performance, others critical thinking, self-regulation, engagement, or assessment. The result is a literature that is directionally informative but heterogeneous. A responsible interpretation therefore avoids statements such as “ChatGPT improves critical thinking” without specifying the learning design and measurement conditions. Instead, the more defensible proposition is conditional: ChatGPT can support critical thinking when instructional design requires students to interrogate, verify, compare, and reflect on AI outputs.

This conditional interpretation is supported by several systematic reviews. Deng et al. (2025) synthesized experimental evidence and reported positive effects of ChatGPT on academic performance, affective-motivational outcomes, and higher-order thinking propensities, while also observing reduced mental effort in some settings. The reduction in mental effort is not automatically negative; it can mean that AI removes low-level burdens and allows students to allocate attention to more complex reasoning. However, it can also represent cognitive offloading when the system performs the central intellectual task. Thus, economics lecturers should distinguish productive offloading from destructive substitution.

The distinction is also visible in experimental work. Urban et al. (2024) found that ChatGPT could improve the quality, elaboration, and originality of university students' creative problem-solving responses, while students perceived tasks as easier and requiring less mental effort. The study therefore illustrates the dual nature of AI support: students can produce better outputs while simultaneously experiencing a lower cognitive burden. Whether that lower burden promotes or undermines learning depends on whether students still perform the critical operations needed to understand and justify the solution.

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

ChatGPT as a Learning Partner

The literature supports several legitimate educational roles for ChatGPT. First, it can act as an on-demand explanatory partner. Students can ask for an alternative explanation, a simpler example, or a step-by-step breakdown. Second, it can support brainstorming by generating multiple possible perspectives before students decide which are worth investigating. Third, it can provide formative feedback on drafts, allowing students to identify unclear reasoning before submitting work. Fourth, it can support simulation and dialogue, such as asking the system to play the role of a consumer, entrepreneur, policy analyst, or central-bank adviser. These uses are pedagogically stronger than asking for a finished assignment because they leave substantive decisions with the student.

Ng et al. (2024) demonstrated the potential of generative AI to support self-regulated learning through planning, monitoring, and feedback activities. Dahri et al. (2024) similarly found strong acceptance of ChatGPT for metacognitive self-regulated learning in a mixed-method study involving preservice teachers. These findings suggest that AI can be incorporated into learning cycles in which students set goals, attempt a task, obtain feedback, monitor progress, and revise. For economics students, such a cycle can be applied to case analysis, policy evaluation, business-plan development, and financial decision-making exercises.

The important pedagogical shift is from answer generation to dialogic learning. A student should be encouraged to ask ChatGPT questions that expose reasoning: “What assumptions does this explanation make?”, “What evidence would falsify this claim?”, “Give me an alternative interpretation”, or “Which part of your answer is uncertain?” Such prompts turn the interaction into a form of intellectual dialogue. Yet even these prompts should not be accepted uncritically; students still need to verify the resulting claims using authoritative sources.

Students' own experiences reinforce the need for structured guidance. Chan and Hu (2023) found generally positive attitudes toward GenAI among university students, while Morell-Mengual et al. (2025) reported widespread chatbot use for understanding complex concepts, time management, and academic performance. At the same time, students reported concerns about dependency and loss of writing or analytical skills. The coexistence of usefulness and concern should be interpreted as a signal for curriculum design: students are already using AI, so higher education should teach them to use it well.

From Productive Support to Cognitive Offloading

Cognitive offloading occurs when an individual relies on an external tool to reduce or replace mental work. Offloading is not inherently harmful; humans routinely use calculators, notes, maps, and search engines. The educational issue arises when the external tool performs the very reasoning that the learner is supposed to develop. In a critical-thinking assignment, for example, asking ChatGPT to identify the strongest argument, select the evidence, and write the conclusion can bypass the student's analysis and evaluation processes.

The literature increasingly identifies this boundary as a central issue. The 2026 systematic review of ChatGPT's cognitive impact found that structured inquiry and scaffolding were associated with positive cognitive development, whereas unstructured contexts could produce cognitive offloading and weaker critical engagement (Deng et al., 2025). The review of critical-thinking skills and dispositions likewise concluded that ChatGPT can support analytical, interpretive, reasoning, and self-regulatory skills when appropriately integrated, but evidence for deeper dispositions such as truth-seeking and systematicity is less consistent (Deng et al., 2025).

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

For economics education, cognitive offloading can be especially problematic because many economic judgments depend on interpretation rather than simple retrieval. Consider a question about whether a government should increase fuel subsidies. ChatGPT can quickly produce arguments for and against. But the educational objective is not to possess a list of arguments. Students need to determine which stakeholders are affected, distinguish short-run from long-run effects, identify fiscal costs, consider price signals, evaluate distributional consequences, examine evidence, and make a conditional policy recommendation. If AI performs these steps, the student may submit a sophisticated answer without developing economic reasoning.

A deep learning approach therefore requires an explicit division of labor between student and AI. AI can generate candidate explanations, counterarguments, examples, and questions. Students should perform verification, prioritization, evidence evaluation, contextual interpretation, and final judgment. This division makes the learning process more demanding but also more educationally meaningful.

Deep Learning as a Pedagogical Filter

Deep learning can be understood as a filter that transforms a generic AI interaction into a cognitively purposeful learning activity. The filter has five components: meaningful context, inquiry, evidence verification, reflection, and independent argumentation. Without these components, AI use may remain at the level of information retrieval. With them, the same AI system can become part of a deeper learning cycle.

The first component is meaningful context. Economic concepts should be connected to authentic problems. Instead of asking students to define inflation, a lecturer might ask them to analyze why food prices in a particular region changed over a given period. The students can use ChatGPT to propose possible mechanisms, but they must then compare these explanations with official statistics, economic reports, and peer-reviewed literature. The local context gives the task a purpose that a generic definition cannot provide.

The second component is inquiry. Students should formulate questions before they seek answers. Inquiry changes the psychological position of the learner from recipient to investigator. A student investigating unemployment might ask: Which unemployment measure is appropriate? What are the limitations of the data? Which groups are most affected? Does the trend indicate cyclical, structural, frictional, or seasonal unemployment? What alternative explanations should be considered? ChatGPT can assist with brainstorming, but the student determines which questions matter.

The third component is evidence verification. AI-generated content is probabilistic and can contain inaccuracies, unsupported claims, or fabricated references. Therefore, students should treat AI outputs as hypotheses requiring verification. In economics, authoritative sources may include national statistical agencies, central banks, government budget documents, international organizations, and peer-reviewed research. A student who learns to triangulate AI claims against these sources develops a form of AI-era information literacy that is directly relevant to professional economic analysis.

The fourth component is reflection. Students should be asked to explain what changed in their understanding after interacting with AI. Reflection can include identifying an AI answer that initially seemed convincing but was later revised, describing which source changed their interpretation, or explaining why two sources reached different conclusions. Such reflection develops metacognitive awareness and makes learning visible.

IJESE
*(*International Journal of Econimics, Science, and Education*)*
Volume 4 Number 6 2026
This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

The fifth component is independent argumentation. The final answer should be the student's reasoned position. AI can be acknowledged as a supporting tool, but the student should own the claim, evidence, assumptions, and limitations. This principle is consistent with current GenAI guidance emphasizing human agency and responsible use in education (UNESCO, 2023).

AI Literacy as a Condition for Critical Thinking

AI literacy is increasingly inseparable from academic literacy. Students need to know not only how to operate a tool but also how the tool produces uncertainty, why outputs can sound confident while being incorrect, and how to evaluate generated claims. Research on AI literacy in higher education suggests that meaningful learning with GenAI depends on students' ability to understand, interact with, evaluate, and regulate AI use (Kohnke et al., 2024; Xie et al., 2025).

The GenAI divide is therefore a significant educational issue. Students do not enter university with equal experience, confidence, access, or critical understanding of AI. The study of university students by Kohnke et al. (2025) shows that students with higher AI literacy were more likely to use GenAI productively, while lower-literacy students were more cautious and fearful. Productive integration therefore requires institutional and course-level AI literacy development rather than assuming that all students naturally know how to use ChatGPT responsibly.

In economics courses, AI literacy can be integrated without creating a separate technical course. Lecturers can teach prompt formulation through economic problems, source verification through policy cases, bias detection through competing economic viewpoints, and hallucination detection through intentionally flawed AI responses. For example, a lecturer may provide three ChatGPT-generated statements about monetary policy and ask students to identify which claims require evidence, which concepts are oversimplified, and which statements are potentially misleading. The activity simultaneously teaches economics and AI literacy.

AI literacy also includes ethical literacy. Students should understand when disclosure of AI use is appropriate, how to protect confidential information, why fabricated citations are problematic, and why submitting AI-generated work as if it were independently produced can undermine academic integrity. Current scholarship increasingly favors transparent, pedagogically justified AI use over either unrestricted use or blanket prohibition (Xia et al., 2024; Yusuf et al., 2024).

Metacognition and Self-Regulated Learning

Critical thinking cannot be separated from metacognition. Students need to monitor whether they understand a concept, recognize when an explanation is insufficient, and decide what evidence is needed next. GenAI creates both opportunities and risks for this process. It can provide immediate feedback and alternative explanations, but it can also create an illusion of understanding because fluent language feels convincing.

Xie et al. (2025) examined metacognitions about GenAI use among 1,229 college students and highlighted the importance of students' beliefs about AI in understanding patterns of use and potential problematic dependence. Yao (2025) similarly identified planning, monitoring, evaluating, information management, and debugging as metacognitive strategies in postgraduate students' use of ChatGPT. These findings reinforce the argument that students should be taught to monitor the interaction rather than simply consume it.

A practical learning cycle for economics can be represented as Plan–Ask–Check–Compare–Explain–Reflect. In the Plan phase, students define the economic problem and identify what they already know. In Ask, they use ChatGPT to generate possible explanations or questions. In Check, they verify AI outputs using credible sources. In Compare, they examine alternative theories or

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

stakeholder perspectives. In Explain, they construct their own argument. In Reflect, they evaluate what they learned and where uncertainty remains. This cycle turns AI use into a metacognitive process.

Assessment Redesign in Economics Education

Assessment is one of the areas most directly disrupted by GenAI. Traditional take-home essays, generic discussion questions, and routine summaries can often be completed with substantial AI assistance. This does not mean that writing assignments are obsolete; rather, assessment needs to make the reasoning process more visible. Reviews of GenAI assessment recommend greater emphasis on self-regulated learning, responsible use, process evidence, and integrity (Xia et al., 2024; Weng et al., 2024; Zhao et al., 2024).

A deep-learning assessment can require students to submit an evidence portfolio. For an economics policy case, the portfolio might include: (1) the student's initial position before AI use; (2) the prompt used with ChatGPT; (3) the AI response; (4) two or three authoritative sources used to verify the response; (5) a table of claims that were accepted, modified, or rejected; (6) the student's final economic argument; and (7) a reflection explaining how the interaction changed the analysis. The assessment then rewards reasoning rather than polished prose alone.

Such an assessment also creates transparency. Students can disclose how AI was used and lecturers can evaluate whether AI supported or replaced the intended learning process. Emerging scholarship on transparency mechanisms indicates that many institutions still lack clear operational procedures for documenting GenAI use, which reinforces the need for explicit course-level expectations (Pérez-Pérez et al., 2026).

The assessment rubric should therefore include dimensions such as conceptual accuracy, evidence quality, source verification, argument coherence, recognition of assumptions, counterargument handling, independent judgment, and responsible AI disclosure. These dimensions align more closely with the goals of economics education than a rubric based primarily on grammatical quality.

Proposed Learning Model: From AI User to Critical Thinker

Table 2 Proposed Learning Model

Stage	Student Behavior	Lecturer Intervention	Evidence of Learning
1. AI User	Requests answers, summaries, or finished text.	Establish boundaries and demonstrate risks.	Prompt/output record.
2. AI-Assisted Learner	Uses AI for explanation, examples, brainstorming, and feedback.	Provide guided prompts and learning objectives.	Learning notes and revised questions.
3. Critical AI User	Verifies claims, checks sources, compares responses, detects errors.	Teach verification, source evaluation, and counterargument.	Verification matrix and source audit.
4. Critical Thinker	Builds independent evidence-based economic arguments and judgments.	Facilitate debate, reflection, and authentic problem solving.	Final argument, evidence portfolio, reflection.

The proposed model suggests that the key developmental transition occurs between the second and third stages. Many students can become AI-Assisted Learners quickly because ChatGPT is easy to access. The more difficult educational task is developing the habits required to question the tool.

IJESE
*(*International Journal of Econimics, Science, and Education*)*
Volume 4 Number 6 2026
This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

Students need repeated practice in identifying unsupported claims, checking sources, comparing explanations, and articulating why they accept or reject an AI-generated proposition.

The final stage is not AI independence in the sense of never using AI. A Critical Thinker can use AI extensively while retaining intellectual responsibility. This distinction is important because professional economists and business practitioners will likely work alongside AI systems. The goal of higher education is therefore not to train students to avoid AI, but to train them to remain capable of judgment when AI is present.

Application Examples in Economics Courses

The framework can be applied across several economics topics. In microeconomics, students can ask ChatGPT to generate competing explanations of a price change and then test the explanations using supply-and-demand theory and actual market information. In macroeconomics, students can compare AI-generated explanations of inflation or unemployment with central-bank and statistical-agency data. In public economics, students can use ChatGPT to generate arguments for and against a tax policy, then evaluate distributional effects and fiscal trade-offs. In entrepreneurship, students can use AI to generate alternative business models but must test assumptions through customer evidence and market analysis. In accounting, students can use AI to explain a transaction but must independently construct the journal entry and justify the accounting treatment.

These examples demonstrate that the deep-learning model is not tied to one course. What changes is the cognitive structure of the assignment. The lecturer moves from "ask and answer" toward "problem, inquiry, evidence, evaluation, argument, reflection." ChatGPT can participate in each stage, but it should never be the sole source of evidence or the final decision-maker.

Table 3 Benefits and Risks: A Balanced Interpretation

Dimension	Potential Benefit	Potential Risk	Deep-Learning Response
Access	Rapid explanations and examples.	Overreliance on generated information.	Require source triangulation.
Efficiency	Reduces routine cognitive load.	Destructive cognitive offloading.	Reserve core reasoning for students.
Feedback	Immediate formative feedback.	Uncritical acceptance.	Require error analysis.
Creativity	Generates alternative ideas.	Homogenized or derivative thinking.	Require original justification.
Personalization	Adapts explanations to student needs.	Unequal AI literacy and access.	Provide common AI-literacy instruction.
Assessment	Supports iterative learning.	Academic integrity problems.	Assess process and disclose AI use.
Metacognition	Supports planning and reflection.	Illusion of understanding.	Require reflective learning logs.

The balanced interpretation is important because neither technological optimism nor technological pessimism is sufficient. The empirical literature contains positive findings, null or mixed findings, and documented risks. The strongest common pattern is the importance of conditions. When AI is embedded in structured inquiry, feedback, verification, and reflection, it can amplify learning. When it is used as an answer substitute, it can weaken the very cognitive processes higher education intends to develop.

IJESE
*(*International Journal of Econimics, Science, and Education*)*
Volume 4 Number 6 2026
This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

Implications for Lecturers and Institutions

At the lecturer level, the first implication is the need to redesign learning tasks. Generic prompts should be replaced with authentic, contextual, and evidence-rich problems. The second is the need to make AI use explicit. Students should know whether AI is permitted, for what purposes, what disclosure is required, and which activities must be completed independently. The third is the need to teach verification. Students should not be expected to infer AI limitations automatically.

At the institutional level, AI policy should balance academic integrity with pedagogical innovation. UNESCO's human-centered guidance emphasizes human agency, ethical safeguards, privacy, and responsible educational use. Institutions should therefore provide staff development, student AI-literacy programs, assessment guidance, and mechanisms for transparent AI disclosure rather than relying only on prohibition.

The literature also indicates that AI integration should not be separated from broader digital inequality. Students differ in access, confidence, prior knowledge, and ability to formulate effective prompts. The GenAI divide identified in recent higher-education research suggests that institutions should provide equitable access and common foundational instruction so that AI becomes a learning resource rather than another source of inequality.

For economics programs, a particularly valuable institutional strategy would be to integrate AI literacy with economic literacy. Students can learn how AI affects markets, labor, productivity, entrepreneurship, and financial decision-making while simultaneously learning how to use AI as an analytical tool. This creates a curriculum that is both economically relevant and digitally future-oriented.

Theoretical Contribution

The article contributes a conditional model of AI-supported learning in which technology is not treated as an independent cause of critical thinking. Instead, deep-learning pedagogy acts as a mediator between AI capability and cognitive development. The model can be expressed as: ChatGPT capability → instructional design → student cognitive activity → evidence evaluation → reflection → critical economic judgment. The implication is that the same AI capability can produce different learning outcomes depending on what the learner is required to do.

This framework extends a simple technology-adoption perspective. Adoption models explain why students use ChatGPT, but adoption alone does not explain whether use is educationally productive. The literature on acceptance, AI literacy, and learning outcomes suggests that perceived usefulness, trust, and facilitating conditions influence use, while pedagogical scaffolding determines what students actually do with the tool (Jo, 2024; Strzelecki, 2025). Deep learning therefore supplies the pedagogical mechanism that connects technology use with intellectual development.

The model also reframes critical thinking in the AI era. Critical thinking now includes not only evaluating human-produced claims but also evaluating machine-generated claims. Students need to recognize plausible-sounding errors, verify citations, inspect assumptions, and compare AI output with evidence. In this sense, AI literacy and critical thinking increasingly overlap. The ability to challenge an AI response becomes a contemporary form of source criticism.

Limitations of the Review

This study has several limitations. First, the evidence base is rapidly changing, so conclusions may need updating as new empirical studies appear. Second, the literature uses heterogeneous definitions and measurements of critical thinking, learning outcomes, AI literacy, and self-regulation. Third, much of the evidence is from general higher education rather than economics education

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

specifically. The proposed application to economics therefore represents a theoretically informed transfer of evidence rather than a claim that all findings have been tested directly with economics students.

Fourth, the PRISMA-oriented structure in this manuscript is intended to provide transparent systematic-review reporting, but the exact database counts should be refreshed by the authors before journal submission if a formal independent Scopus/WoS export is performed. This is preferable to presenting unstable database counts as permanent facts. Fifth, the review focuses primarily on ChatGPT and closely related GenAI literature; future work should compare multiple models and investigate whether differences in model capability influence learning outcomes.

Finally, the review does not estimate an overall effect size. A future meta-analysis could examine whether structured ChatGPT interventions produce larger effects on critical thinking than unstructured use and whether outcomes differ by discipline, task type, AI literacy, duration, and assessment method.

D. CONCLUSION

ChatGPT has changed the conditions under which university students learn economics. Its ability to provide explanations, generate alternatives, offer feedback, and support problem solving creates substantial opportunities for learning. However, the same capability can produce dependency and cognitive offloading when students use AI primarily to obtain finished answers. The key educational issue is therefore not whether students use ChatGPT, but what they are required to think about while using it.

The systematic synthesis indicates that deep learning provides a useful pedagogical response. Deep learning requires meaningful context, inquiry, evidence verification, reflection, and independent argumentation. These elements transform ChatGPT from an answer-producing device into a learning partner whose outputs become objects of analysis. The proposed progression from AI User to AI-Assisted Learner, Critical AI User, and finally Critical Thinker provides a practical framework for designing economics learning activities.

For economics education, the model is particularly relevant because economic problems involve evidence, uncertainty, assumptions, competing interests, and policy trade-offs. Students should therefore be assessed not only on the final answer but also on how they formulate questions, verify claims, evaluate evidence, handle counterarguments, and justify conclusions. AI literacy should be integrated with economic literacy so that students can use GenAI responsibly while maintaining intellectual agency.

The central conclusion is that ChatGPT should not replace the student's thinking process. It should make that process more visible, more questioning, and more evidence-oriented. The future-ready economics student is not a person who refuses AI, nor a person who accepts AI uncritically, but a critical thinker who knows when to use AI, how to challenge it, and how to make an independent judgment when the evidence remains uncertain.

Future empirical studies should test the proposed model with economics students using experimental or quasi-experimental designs. Research can compare conventional ChatGPT use with structured deep-learning interventions and measure critical thinking, economic literacy, metacognition, source verification, and academic integrity. Such studies would provide stronger causal evidence about whether the transition from AI User to Critical Thinker can be systematically achieved.

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

E. REFERENCES

Abdallah, N., Katmah, R., Khalaf, K., & Jelinek, H. F. (2025). Systematic review of ChatGPT in higher education: Navigating impact on learning, wellbeing, and collaboration. Social Sciences & Humanities Open, 12, 101866. https://doi.org/10.1016/j.ssaho.2025.101866

Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in the 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

Bhullar, P. S., Joshi, M., & Chugh, R. (2024). ChatGPT in higher education—A synthesis of the literature and a future research agenda. Education and Information Technologies, 29, 21501–21522. https://doi.org/10.1007/s10639-024-12723-x

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in Higher Education, 20, 43. https://doi.org/10.1186/s41239-023-00411-8

Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. Educational Technology Research and Development, 72, 2401–2416. https://doi.org/10.1007/s11423-024-10366-w

Dahri, N. A., Yahaya, N., & Al-Rahmi, W. M. (2024). Extended TAM based acceptance of AI-powered ChatGPT for supporting metacognitive self-regulated learning in education: A mixed-methods study. Heliyon, 10(8), e29317. https://doi.org/10.1016/j.heliyon.2024.e29317

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

Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? Journal of Responsible Technology, 13, 100060. https://doi.org/10.1016/j.jrt.2023.100060

Jo, H. (2024). From concerns to benefits: A comprehensive study of ChatGPT usage in education. International Journal of Educational Technology in Higher Education, 21, 35. https://doi.org/10.1186/s41239-024-00471-4

Jensen, L. X., Buhl, A., Sharma, A., & Bearman, M. (2025). Generative AI and higher education: A review of claims from the first months of ChatGPT. Higher Education, 89, 1145–1161. https://doi.org/10.1007/s10734-024-01265-3

Kohnke, L., Zou, D., & Zhang, R. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. Computers and Education Open, 6, 100184. https://doi.org/10.1016/j.caeo.2024.100184

Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. International Journal of Educational Technology in Higher Education, 20, 56. https://doi.org/10.1186/s41239-023-00426-1

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

Lo, C. K., Hew, K. F., & Jong, M. S.-Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. Computers & Education, 219, 105100. https://doi.org/10.1016/j.compedu.2024.105100

McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025). Generative AI chatbots in higher education: A review of an emerging research area. Higher Education, 89, 1533–1549. https://doi.org/10.1007/s10734-024-01288-w

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. Educational Process: International Journal, 14, e2025031. https://doi.org/10.22521/edupij.2025.14.31

Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J. M., & López-Meneses, E. (2023). Impact of the implementation of ChatGPT in education: A systematic review. Computers, 12(8), 153. https://doi.org/10.3390/computers12080153

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. British Journal of Educational Technology, 55, 1328–1353. https://doi.org/10.1111/bjet.13454

Ngo Cong-Lem, & Nguyen Thi Thuy-Dung. (2025). Towards a framework for understanding and assessing critical thinking in generative AI-enhanced learning environments: A scoping review. BehaviourWorks Australia.

This work is licensed under a Creative Commons Attribution- ShareAlike 4.0 International License.
ISSN : 3046-689X (ONLINE)

Pérez-Pérez, I., González-Afonso, M. C., Plasencia-Carballo, Z., & Pérez-Jorge, D. (2026). Transparency mechanisms for generative AI use in higher education assessment: A systematic scoping review (2022–2026). Computers, 15(2), 111. https://doi.org/10.3390/computers15020111

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

Strzelecki, A. (2025). ChatGPT in higher education: Investigating bachelor and master students' expectations towards AI tool. Education and Information Technologies, 30, 10231–10255. https://doi.org/10.1007/s10639-024-13222-9

Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. Computers & Education, 215, 105031. https://doi.org/10.1016/j.compedu.2024.105031

UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.

Weng, X., Xia, Q., Gu, M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. Australasian Journal of Educational Technology, 40(6), 37–55. https://doi.org/10.14742/ajet.9540

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. International Journal of Educational Technology in Higher Education, 21, 40. https://doi.org/10.1186/s41239-024-00468-z

Xie, Y., Li, Y., Yu, T., & Liu, Y. (2025). Metacognitions about generative AI use: Psychometric and network analysis among Chinese college students. Education and Information Technologies, 30, 20523–20542. https://doi.org/10.1007/s10639-025-13584-8