BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg
BG PULSE JOURNAL

GENERATIVE AI AND HUMAN THINKING: COGNITIVE OFFLOADING, LEARNING, MEMORY, AND CRITICAL THINKING

Jurahanova Nigora
Kimyo International University in Tashkent, Namangan branch
Doi: https://doi.org/10.5281/zenodo.23754343

Abstract. Conversational artificial intelligence (AI) systems such as ChatGPT and Claude have rapidly become everyday tools for hundreds of millions of people. Unlike earlier tools, they do not merely retrieve information but analyse, write and draw conclusions on the user's behalf. This narrative review analyses the available empirical evidence on how offloading cognitive tasks to generative AI affects brain engagement, learning, memory and critical thinking. It focuses on an EEG study from the MIT Media Lab (n=54), a survey of 319 knowledge workers by Microsoft Research and Carnegie Mellon University, a randomised field experiment with high-school students and related research. The evidence indicates that uncritical reliance on AI is associated with lower engagement of the user's own cognitive resources during the task, poorer recall of self-written text and a reduced tendency to verify AI output; the higher the trust in AI, the less critical scrutiny is applied. However, the evidence base is still limited: studies are small, some are non-peer-reviewed preprints, and no long-term harm has been demonstrated. A conceptual model of "cognitive debt" accumulation and recommendations for the reasoned use of AI in higher education, including medical education, are proposed.

Keywords: generative artificial intelligence, cognitive offloading, cognitive debt, critical thinking, EEG, learning, higher education, medical education.

Introduction

People have always used external tools to reduce mental effort — writing, calculators, notebooks and search engines. In cognitive psychology this is described as cognitive offloading: the use of physical action or external resources to reduce the cognitive demands of a task [18]. The theory of the extended mind goes further and argues that, under certain conditions, external resources can function as part of the cognitive system itself [5].

Generative AI changes the nature of this offloading. A calculator returns a number that the user must interpret; a search engine returns sources that the user must read, compare and synthesise. A large language model can perform the interpretation, comparison and synthesis itself and deliver a finished answer, argument or text. Several distinct thinking steps that previously required independent engagement are thereby collapsed into a single step that does not.

The adoption of these tools in education has been extremely rapid. Large language models offer opportunities for personalised explanation and feedback, but also raise concerns about over-reliance, reduced independent thinking and academic integrity [12]. Uzbekistan has adopted a national Strategy for the Development of Artificial Intelligence Technologies until 2030, which envisages the wide introduction of AI, including in education [25]. This makes it important to understand the cognitive consequences of regular AI use for students and teachers.

Since 2024, the first experimental, neurophysiological and survey data on the cognitive effects of generative AI use have appeared [13; 15; 22]. Some of these studies attracted wide public

attention and were often reported in simplified form ("ChatGPT damages the brain"). A careful, balanced analysis of what this evidence does and does not show is therefore needed.

The aim of this article is to analyse the empirical evidence on the effects of offloading cognitive tasks to generative AI on brain engagement, learning and critical thinking, to propose a conceptual model of the underlying mechanism, and to formulate recommendations for higher education.

The article addresses three questions:

1. What do neurophysiological and behavioural studies show about the effects of regular use of generative AI on cognitive engagement?
2. Through which mechanisms can AI reliance affect learning and critical thinking?
3. How can AI be used in education without accumulating "cognitive debt"?

Materials and methods

The study is designed as a narrative review. Sources were searched in arXiv, the ACM Digital Library, PubMed, Scopus and Google Scholar using combinations of the terms "ChatGPT", "large language model", "generative AI", "cognitive offloading", "critical thinking", "EEG", "learning", "memory" and "over-reliance". Studies published between 2023 and 2026 that reported original empirical data (neurophysiological, experimental or survey) on the cognitive effects of generative AI use were included. Foundational works on cognitive offloading, memory, learning and automation bias were added to interpret the findings.

Given the novelty of the field, preprints were included if they were the primary source of a widely discussed finding; their status is explicitly indicated, and published critical commentaries on them were also considered. No original data were collected.

Results

1. Neurophysiological evidence: the MIT Media Lab EEG study

The most direct evidence comes from a study by Kosmyna et al. at the MIT Media Lab [13]. Fifty-four participants were divided into three groups and wrote essays over three sessions: one group used ChatGPT (LLM group), one used a conventional search engine (Search Engine group), and one used no external tools (Brain-only group). Brain activity was recorded using 32-channel electroencephalography (EEG), and neural connectivity was analysed across frequency bands.

The main findings were as follows. The Brain-only group showed the strongest and most widely distributed neural connectivity; the Search Engine group showed intermediate engagement; and the LLM group showed the weakest connectivity. Over successive sessions, participants in the LLM group increasingly relied on copying generated text, and their essays became more homogeneous in vocabulary and structure. In the first session, 83.3% of LLM users were unable to correctly quote a sentence from the essay they had written minutes earlier, compared with 11.1% in the Brain-only group, and LLM users reported the lowest sense of ownership of their essays [13].

In a fourth, cross-over session (n=18), participants from the LLM group wrote without tools (LLM-to-Brain), while participants from the Brain-only group used ChatGPT for the first time (Brain-to-LLM). The LLM-to-Brain participants showed weaker neural connectivity and under-engagement of alpha and beta networks. In contrast, Brain-to-LLM participants showed higher memory recall and re-engagement of widespread brain regions, and used the AI more to seek information than to obtain a finished text. The authors introduced the term "cognitive debt" to

describe a situation in which reliance on AI saves mental effort in the short term but may create long-term costs, including reduced critical engagement and creativity [13].

The study has important limitations. The sample was small, only 18 participants completed the cross-over session, participants were drawn from a limited population of university students and staff, and the work remains a preprint without completed peer review. A published commentary by Stanković et al. calls for a more cautious interpretation of its design, statistical analysis and EEG methodology [23]. In addition, scalp EEG cannot directly measure activity in deep structures involved in memory, such as the hippocampus.

2. Behavioural evidence: trust in AI and critical thinking

Lee et al. (Microsoft Research and Carnegie Mellon University) surveyed 319 knowledge workers who described 936 real examples of using generative AI in their work [15]. The study was presented at the CHI 2025 conference on human factors in computing systems. It measured self-reported critical thinking and cognitive effort rather than brain activity.

Two opposing effects were identified. Higher confidence in the AI's ability to perform a task was associated with less critical thinking, whereas higher self-confidence in one's own ability to perform the task was associated with more critical thinking. Respondents also reported that generative AI shifted the nature of their critical thinking from information gathering towards information verification, from problem solving towards integrating AI responses, and from task execution towards task stewardship. Barriers to critical engagement included lack of awareness, lack of motivation and lack of ability to inspect or improve AI output [15]. Lack of ability is thus itself a barrier: those least able to evaluate AI output may be the least likely to scrutinise it, although this is an interpretation of self-reported data rather than a directly tested result.

Similar associations were reported by Gerlich, who surveyed 666 participants and found a significant negative correlation between frequent AI tool use and critical thinking abilities, mediated by increased cognitive offloading; younger participants showed higher dependence on AI tools and lower critical thinking scores [10]. As this study is cross-sectional, it cannot establish causality.

A further behavioural finding concerns metacognition, i.e. the accuracy of people's judgements about their own performance. In two studies (n=246 and n=452), Fernandes et al. asked participants to solve Law School Admission Test logical-reasoning problems with the help of ChatGPT [9]. Participants scored about three points higher than a norm population but overestimated their own score by about four points; higher AI literacy was associated with lower, not higher, metacognitive accuracy, and the typical Dunning–Kruger pattern [14] (the least competent overestimating most) was no longer observed with AI use [9]. AI use may thus improve the product while weakening the user's ability to judge how much of it they actually understand.

3. Experimental evidence on learning outcomes

The strongest causal evidence on learning comes from a randomised field experiment by Bastani et al. involving nearly one thousand high-school students in Turkey [2]. Students practised mathematics with access to either a standard GPT-4 interface (GPT Base), a version designed with teacher-input safeguards that provided hints instead of answers (GPT Tutor), or no AI. During practice, both AI groups performed substantially better. However, on a subsequent exam without AI, students who had used GPT Base performed worse than the control group, whereas the negative

effect was largely mitigated in the GPT Tutor group. The authors concluded that students used the unrestricted tool as a "crutch", obtaining answers without learning the underlying skills [2].

This finding is consistent with long-established principles of learning psychology: information that a person generates independently is remembered better than information that is merely read (the generation effect) [20], and active retrieval during learning strengthens long-term retention (the testing effect) [19]. When AI performs the generation and retrieval, these mechanisms are bypassed.

Several smaller experiments point in the same direction. Stadler et al. randomly assigned 91 university students to research a scientific question with ChatGPT or with Google; the ChatGPT group reported lower cognitive load but produced lower-quality reasoning and argumentation in their final recommendations [22]. Fan et al. compared 117 students writing with ChatGPT, a human expert, writing-analytics tools or no additional support: the ChatGPT group achieved larger improvements in essay scores, but knowledge gain and knowledge transfer did not differ significantly, a pattern the authors described as "metacognitive laziness" [8]. Melumad and Yun, in a series of experiments with participants recruited online (n=1,104 and n=1,979 in the first two), found that people who learned about a topic from LLM summaries rather than from web search spent less time, reported learning less, and wrote advice that was sparser and less original [16]. A systematic review of 14 studies in higher education likewise concluded that over-reliance on AI dialogue systems may weaken critical and analytical thinking and decision-making, while also noting benefits for writing efficiency and self-confidence [24].

4. Effects on creativity and diversity of ideas

Doshi and Hauser found in an experiment on short story writing that access to AI-generated ideas made individual stories more creative and better written, especially for less creative writers, but made stories by different authors more similar to each other [7]. This reduction in collective diversity mirrors the homogeneity of essays observed in the LLM group of the MIT study [13].

5. Historical analogues: earlier tools

Concerns about cognitive offloading are not new. Sparrow et al. showed that when people expect information to remain accessible on a computer, they remember the information itself less well, but remember better where to find it [21]. Barr et al. found that people with a more intuitive cognitive style relied more on smartphones to search for information [1]. Dahmani and Bohbot reported that greater lifetime use of GPS navigation was associated with poorer spatial memory, and that greater use over a three-year period was associated with a steeper decline in hippocampus-dependent spatial memory [6]. These studies suggest that cognitive functions that are regularly delegated to external tools may be exercised less, although the magnitude and reversibility of such effects remain uncertain.

6. Evidence from clinical practice and medical decision-making

Two studies from medicine are particularly relevant for medical education. In a randomised clinical trial, Goh et al. gave 50 physicians (attendings and residents) clinical vignettes to solve with either GPT-4 plus conventional resources or conventional resources alone; access to the LLM did not significantly improve diagnostic reasoning (median score 76% versus 74%; adjusted

difference 2 percentage points, 95% CI −4 to 8), whereas the LLM on its own scored 16 percentage points higher than physicians using conventional resources [11]. This shows that having a capable tool does not automatically translate into better human reasoning. In a multicentre observational study in Poland, Budzyń et al. found that the adenoma detection rate in colonoscopies performed without AI assistance fell from 28.4% before to 22.4% after AI polyp-detection systems were introduced in the participating centres (absolute difference −6.0 percentage points; adjusted OR 0.69, 95% CI 0.53–0.89) [4]. The authors interpret this as possible "deskilling"; because the design is observational and based on four centres, causality cannot be established, but the result illustrates the type of risk that would be most serious in clinical training.

Table 1 summarises the key studies.

Table 1. Key empirical studies on cognitive effects of AI and digital tools

Study	Design and sample	Main finding
Kosmyna et al., 2025 [13]	EEG, 3 groups, n=54 (cross-over n=18); preprint	Weakest connectivity and poorest recall of own essay in LLM group
Lee et al., 2025 [15]	Survey, 319 workers, 936 cases	Higher trust in AI — less critical thinking; higher self-confidence — more
Gerlich, 2025 [10]	Survey, n=666	Negative correlation between AI use and critical thinking, mediated by offloading
Bastani et al., 2025 [2]	Randomised field experiment, ≈1,000 students	Unrestricted GPT-4 improved practice but worsened exam results; tutor version mitigated harm
Doshi & Hauser, 2024 [7]	Online experiment, story writing	AI improved individual creativity but reduced collective diversity
Sparrow et al., 2011 [21]	Laboratory experiments	Expecting future access reduces memory for content
Dahmani & Bohbot, 2020 [6]	Cross-sectional and 3-year follow-up	GPS use associated with poorer spatial memory
Stadler et al., 2024 [22]	Randomised experiment, 91 students, ChatGPT vs Google	Lower cognitive load but lower-quality reasoning with ChatGPT
Fan et al., 2025 [8]	Randomised experiment, 117 students	Higher essay scores with ChatGPT, but no difference in knowledge gain or transfer ("metacognitive laziness")

Study	Design and sample	Main finding
Melumad & Yun, 2025 [16]	Online experiments (n=1,104 and n=1,979 in Exp. 1–2)	Learning from LLM summaries produced shallower knowledge than web search
Fernandes et al., 2026 [9]	Two studies, n=246 and n=452	AI use raised performance but users overestimated it; higher AI literacy, lower metacognitive accuracy
Goh et al., 2024 [11]	Randomised clinical trial, 50 physicians	LLM access did not significantly improve diagnostic reasoning
Budzyń et al., 2025 [4]	Multicentre observational study, 1,443 colonoscopies	Adenoma detection without AI fell from 28.4% to 22.4% after AI exposure

Source: compiled by the author.

Discussion

1. A conceptual model of "cognitive debt"

The studies reviewed use very different methods — EEG in the laboratory, surveys of professionals, a field experiment in schools — yet they point in the same direction. Together they suggest a mechanism that can be described as a self-reinforcing cycle (Table 2).

Table 2. Conceptual model of cognitive debt accumulation

Stage	Process	Evidence and theoretical basis
1	Delegation of a cognitive task (writing, analysis, problem solving) to AI	Cognitive offloading [18]
2	Reduced engagement of one's own cognitive resources during the task	Weaker EEG connectivity [13]
3	Weaker encoding and learning	Poor recall of own essay [13]; poorer exam results [2]; generation and testing effects [19; 20]
4	Increased trust in AI and reduced verification	Trust–critical thinking relationship [15]; automation complacency [17]

Stage	Process	Evidence and theoretical basis
5	Reduced competence to evaluate AI output, leading to further delegation	Lack of ability as barrier to critical thinking [15]; overconfidence of the less competent [14]; AI users overestimate their own performance [9]

Source: developed by the author.

The model resembles the well-known phenomenon of automation complacency, in which operators of highly reliable automated systems monitor them less carefully and fail to detect occasional errors [17]. In the case of generative AI, the risk is amplified by two features: AI output is fluent and persuasive, and AI systems produce errors (including fabricated facts and references) in a way that is not obvious to the user [15].

The model also implies that the effect is not inevitable. The key variable is not the tool itself but how it is used. In the MIT study, participants who first did the work independently and then used the AI engaged differently and showed better recall [13]; in the Bastani experiment, a tutor version that gave hints instead of answers largely removed the harm [2]. This is consistent with research showing that effortful processing, such as active retrieval and self-generation, strengthens learning, whereas tools that eliminate all effort may undermine it [19; 20].

Human–computer interaction research offers practical tools for reducing this risk. Buçinca et al. showed in an experiment with 199 participants that "cognitive forcing functions" — interface designs that make the user think before or while receiving the AI's suggestion, for example by committing to an answer first or requesting AI advice only on demand — reduced overreliance on incorrect AI suggestions more than simply adding explanations; however, the designs that were most effective were also rated least favourably by users, and the benefit was larger for people with a higher need for cognition [3]. This suggests that safeguards against cognitive debt may need to be built into the learning environment rather than left to the learner's willingness.

2. Implications for higher and medical education

The greatest risks arise in education, where the goal of a task is not the product (an essay, a solved problem) but the cognitive process it is designed to develop. A student who submits an AI-generated essay may complete the assignment without performing the retrieval, structuring and self-correction that the assignment was intended to train. The MIT authors warn that this may be particularly harmful for younger users whose reasoning and writing skills are still developing [13].

In medical education, the stakes are higher. Clinical reasoning requires the ability to independently generate differential diagnoses, evaluate evidence and recognise errors, including errors in AI-generated recommendations. Over-reliance on AI during training may reduce opportunities to develop exactly these skills. Evidence from clinical settings is consistent with this concern: access to an LLM did not by itself improve physicians' diagnostic reasoning [11], and routine AI assistance in colonoscopy was followed by a lower detection rate when clinicians worked without it [4]. At the same time, AI can be valuable for explanation, feedback and simulation if it is used to support rather than replace independent thinking.

Based on the evidence reviewed, the following principles for educational practice are proposed (Table 3).

Table 3. Principles for using generative AI without accumulating cognitive debt

Principle	Practical implementation
Think first, then ask AI	Write an outline, attempt the problem or form an opinion before consulting AI; use AI to refine or challenge the result
Treat AI output as a claim	Verify key statements against reliable sources, especially in tasks where AI seems to "just know" the answer
Require reasoning	Ask AI to explain its reasoning and sources, and check them independently
Use AI as a tutor, not an answer key	In learning tasks, configure or prompt AI to give hints and feedback rather than complete solutions
Practise without AI regularly	Periodically complete tasks unaided to maintain independent skills
Assess the process	Use oral defence, in-class writing and reflective commentary in addition to written products
Build in friction	Where possible, require the learner to commit to an answer before the AI response is shown, or make AI advice available only on request (cognitive forcing) [3]

Source: developed by the author based on [2; 3; 13; 15; 19].

3. Replacement versus support: mechanisms linking AI use and learning

The findings reviewed in this article should not be interpreted as evidence that generative AI inevitably weakens cognition. A useful distinction is between replacement and support. Replacement occurs when a student delegates the central intellectual work—generating an explanation, solving a problem, selecting evidence, or forming a conclusion—to the system. Support occurs when the student first attempts the task and then uses AI to identify gaps, compare alternatives, or receive feedback. These approaches can produce very different learning experiences.

Cognitive engagement and memory

Learning depends not only on exposure to correct information but also on the mental operations performed by the learner. Recalling information, explaining it in one's own words, making predictions, and correcting mistakes help strengthen accessible knowledge. If a student only reads a polished AI answer, the answer may feel familiar without being reliably retrievable later. This distinction between recognition and independent recall is particularly important before examinations and in professional training.

Verification and automation bias

Fluent, confident language can make an AI-generated answer appear more reliable than it is. Generative systems may omit important qualifications, combine facts incorrectly, or produce references that do not support the stated claim. Verification should therefore be an explicit step. For academic work, students should open cited sources, confirm that they exist, and check whether they support the statement. In clinical learning, AI output must not replace textbooks, current clinical guidance, or the judgement of qualified professionals.

Relevance to medical and dental education

Medical and dental students need more than the ability to obtain an answer: they must learn to explain mechanisms, distinguish similar conditions, assess evidence, and recognise uncertainty. AI can be used productively to generate practice questions, explain a difficult concept at different levels, simulate an oral examination, or critique a student's provisional answer. A safer sequence is to formulate an initial answer independently, request feedback or hints, verify key claims in authoritative sources, and then explain the final answer without assistance. Tasks involving diagnosis, treatment, medication, or patient safety require particular caution and supervision.

A practical protocol for students

A four-step protocol may help preserve independent thinking: (1) Attempt—write an initial answer or solution before consulting AI. (2) Question—ask the system to explain, give hints, or identify weaknesses rather than immediately provide a finished response. (3) Verify—check important claims against reliable sources. (4) Retrieve—close the AI conversation and reproduce the main explanation from memory. If the student cannot explain the result independently, the task should be treated as unfinished learning rather than completed work.

Institutional policy

Educational institutions should make expectations about AI use explicit, including when assistance is permitted, how it should be acknowledged, and which tasks must be completed independently. Assessment can combine written assignments with oral explanation, supervised problem-solving, and short reflections describing how AI was used and how its output was checked.

4. What the evidence does not show

It is important to state clearly the limits of the current evidence. None of the studies reviewed shows that the use of ChatGPT, Claude or similar tools causes permanent brain damage or an irreversible decline in intelligence. The MIT study is small, has not completed peer review and has been critically commented on [13; 23]; EEG reflects cortical activity during a task rather than structural changes. The studies by Lee et al. and Gerlich are based on self-reports and cross-sectional data and cannot establish causality [10; 15]. The Bastani experiment is the most rigorous, but it concerns mathematics in one school system [2]. Evidence from newer experiments is also limited: the study by Fan et al. had 117 participants [8], the experiments by Melumad and Yun used paid online participants and in part hypothetical scenarios [16], and the colonoscopy study is observational and involved four centres [4]. What these studies jointly support is a more modest conclusion: habitual, uncritical delegation of thinking tasks to AI is associated with reduced engagement of one's own cognitive resources during those tasks and with reduced checking of results.

Limitations of this review. This is a narrative rather than a systematic review; the field is developing rapidly and new studies may change the conclusions. Some key sources are preprints. The proposed model is conceptual and requires empirical testing.

Directions for future research include longitudinal studies of cognitive effects of AI use; replication of neurophysiological studies with larger and more diverse samples; randomised trials comparing different modes of AI use (answer-giving versus tutoring); and studies among students in Uzbekistan, including medical students, measuring AI use patterns, critical thinking and learning outcomes. Future studies should also distinguish the frequency of AI use from its purpose and quality.

Conclusion

1. Generative AI differs from earlier tools in the completeness of cognitive delegation: it can perform interpretation, synthesis and writing on the user's behalf.
2. Early empirical evidence indicates that uncritical reliance on AI is associated with reduced brain engagement during the task, poorer memory for one's own work, reduced checking of AI output and, in learning contexts, poorer independent performance.
3. Trust is a key factor: the more a person trusts AI and the less confident they are in their own competence, the less critically they evaluate AI output; users may also overestimate their own performance when working with AI [9].
4. The effects depend on how AI is used rather than on the tool itself. Independent work before using AI, verification of AI output and tutor-style use that provides hints rather than answers reduce the risks.
5. The evidence base remains limited, and claims of lasting brain damage are not supported. Universities, including medical universities in Uzbekistan, should develop AI-use policies and teaching practices that preserve independent thinking, and should support local empirical research on this topic.

References / Adabiyotlar

1. Barr, N., Pennycook, G., Stolz, J. A., & Fugelsang, J. A. (2015). The brain in your pocket: Evidence that smartphones are used to supplant thinking. *Computers in Human Behavior*, 48, 473–480. https://doi.org/10.1016/j.chb.2015.02.029
2. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122, e2422633122. https://doi.org/10.1073/pnas.2422633122
3. Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction, 5*(CSCW1). https://doi.org/10.1145/3449287
4. Budzyń, K., Romańczyk, M., Kitala, D., Kołodziej, P., Bugajski, M., Adami, H. O., Blom, J., Buszkiewicz, M., Halvorsen, N., Hassan, C., et al. (2025). Endoscopist deskilling after exposure to artificial intelligence in colonoscopy: A multicenter observational study. *The Lancet Gastroenterology & Hepatology, 10*(10), 896–903. https://doi.org/10.1016/S2468-1253(25)00133-5
5. Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis*, 58(1), 7–19. https://doi.org/10.1093/analys/58.1.7

6. Dahmani, L., & Bohbot, V. D. (2020). Habitual use of GPS negatively impacts spatial memory during self-guided navigation. *Scientific Reports*, 10, 6310. https://doi.org/10.1038/s41598-020-62877-0
7. Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28), eadn5290. https://doi.org/10.1126/sciadv.adn5290
8. Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13544
9. Fernandes, D., Villa, S., Nicholls, S., Haavisto, O., Buschek, D., Schmidt, A., Kosch, T., Shen, C., & Welsch, R. (2026). AI makes you smarter but none the wiser: The disconnect between performance and metacognition. *Computers in Human Behavior, 175*, 108779. https://doi.org/10.1016/j.chb.2025.108779
10. Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. https://doi.org/10.3390/soc15010006
11. Goh, E., Gallo, R., Hom, J., Strong, E., Weng, Y., Kerman, H., et al. (2024). Large language model influence on diagnostic reasoning: A randomized clinical trial. *JAMA Network Open, 7*(10), e2440969. https://doi.org/10.1001/jamanetworkopen.2024.40969
12. 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
13. Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2506.08872
14. Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121
15. Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3706598.3713778
16. Melumad, S., & Yun, J. H. (2025). Experimental evidence of the effects of large language models versus web search on depth of learning. *PNAS Nexus*, pgaf316. https://doi.org/10.1093/pnasnexus/pgaf316
17. Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410. https://doi.org/10.1177/0018720810376055
18. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002
19. Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

20. Slamecka, N. J., & Graf, P. (1978). The generation effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory*, 4(6), 592–604. https://doi.org/10.1037/0278-7393.4.6.592
21. Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, 333(6043), 776–778. https://doi.org/10.1126/science.1207745
22. Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior, 160*, 108386. https://doi.org/10.1016/j.chb.2024.108386
23. Stanković, M., Hirche, E., Kollatzsch, S., & Doetsch, J. N. (2026). *Comment on: Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2601.00856
24. Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*, 28. https://doi.org/10.1186/s40561-024-00316-7
25. O'zbekiston Respublikasi Prezidentining 2024-yil 14-oktabrdagi PQ-358-son "2030-yilgacha sun'iy intellekt texnologiyalarini rivojlantirish strategiyasini tasdiqlash to'g'risida"gi qarori. https://lex.uz/docs/7159258