TECHNOLOGY

How to Use AI Tools to Write, Study and Research Faster: An Evidence-Based Beginner's Guide

Ateendra Jha*, Meena Gupta, Yumi Nakamura

* *Correspondence: Ateendra Jha*

Received: 26 November 2026 **Published:** 26 November 2026 **Volume 13, Issue 4** (Oct–Dec 2026)
Article no.: e240 **Open access** · free to read

ABSTRACT

Generative artificial intelligence (AI) tools, especially large language model (LLM) chatbots, have moved from novelty to routine in education and knowledge work. In 2025–2026 surveys of UK undergraduates, 92–95% reported using AI in some form.[1,2] This beginner's guide reviews experimental and survey evidence on whether these tools help people write, study and research faster, and when they do not. In a preregistered experiment, ChatGPT reduced the time taken on professional writing tasks by about 40% while raising rated quality by 18%.[3] However, a field experiment found that unrestricted AI access improved practice performance but later reduced exam-style performance without AI,? a small preprint reported weaker neural engagement and poorer recall of one's own essays with LLM use,? and seven experiments found shallower learning from LLM summaries than from web search.? Chatbots can also fabricate references,? and AI-text detectors misclassify non-native English writing.? A widely cited meta-analysis reporting large learning gains from ChatGPT was retracted in 2026.?,[1]? The article proposes a practical Draft–Check–Own workflow: use AI to accelerate drafting, explanation and practice, but verify every fact and reference, and do the core thinking yourself. Speed gains in writing are well supported; learning benefits are conditional on tool design and user behaviour.

Keywords: *generative AI, large language models, AI-assisted writing, study skills, literature search, hallucination, cognitive offloading, AI literacy*

Introduction

AI use among students has become close to universal. In a December 2025 survey of 1,054 full-time UK undergraduates, 95% reported using AI in at least one way and 94% used generative AI to help with assessed work.[1] The previous edition, based on 1,041 students, found 92% and 88%, up from 53% using generative AI for assessments in 2024; in that edition only 36% of students said their institution had trained them in AI skills.[2]

Beginners therefore receive two conflicting messages: that AI saves large amounts of time, and that it weakens thinking and invents facts. Both messages have experimental support, and they are not contradictory once the *task* and the *way of using the tool* are separated. This article asks: for writing, studying and research, which uses of AI tools are supported by evidence, which are risky, and what simple workflow follows from that evidence? It is written for students, early-career researchers and professionals who are new to these tools.

Scope and Evidence Base

This is a narrative evidence review, not a systematic review. It draws on peer-reviewed experiments, one national survey series, one classic learning-science study, and one preprint that has not been peer reviewed (flagged where used). Three limitations matter. First, most experiments tested earlier models (such as GPT-3.5 and GPT-4), and current tools change quickly, so numerical effects may not transfer. Second, the survey data are from the United Kingdom; no equivalent India-specific national dataset was verified in this review, so none is reported. Third, one influential meta-analysis was found to have been retracted during research for this article, and it is therefore discussed as a cautionary example rather than as evidence.

1. What AI Writing and Study Tools Actually Do

LLMs generate fluent text by predicting likely continuations of a prompt. This makes them useful for drafting, rewording, explaining and generating practice material, but it also means that output can sound authoritative while being wrong. Researchers who audited ChatGPT-generated references stress that such systems produce plausible text rather than retrieve verified records.? For beginners, it helps to sort uses into four groups:

• **Generate:** outlines, first drafts, alternative phrasings.

• **Transform:** summarising, simplifying, translating or changing tone.

• **Explain and tutor:** step-by-step explanations, quizzes, feedback on answers.

• **Retrieve:** tools that search and cite sources; these still require checking.

Each group has a different risk profile, discussed below.

2. Writing Faster: What the Evidence Shows

The strongest evidence for speed comes from writing tasks. Noy and Zhang assigned occupation-specific, incentivised writing tasks to 453 college-educated professionals and randomly exposed half to ChatGPT. Average completion time fell by 40% and rated output quality rose by 18%; inequality in performance between workers decreased; and exposed participants were twice as likely to report using ChatGPT in their real job two weeks later.[3]

Three cautions apply. The tasks were short, online and based on GPT-3.5.[3] The authors themselves noted that real-world speed gains are likely smaller because users must write prompts and fact-check output.[11] And these results concern *producing a deliverable*, not *learning to write*, which is a different goal (see Section 3).

Practical implication (author's synthesis): use AI for outlines, alternative wordings, tone adjustment and feedback on clarity. Write your own central argument first, then edit AI suggestions critically rather than accepting them wholesale.

3. Studying Faster Without Learning Less

3.1 How students use AI

In the 2025 HEPI survey, explaining concepts was the most common use of generative AI for assessments (58% of students), alongside summarising articles and suggesting research ideas.[2] These are study-support uses, which is where the learning evidence becomes important.

3.2 Performance with AI is not the same as learning

Bastani and colleagues ran a field experiment with nearly 1,000 high school mathematics students using two GPT-4 tutors: a standard chat interface ("GPT Base") and a version with prompts designed to safeguard learning ("GPT Tutor"). Students with AI access did better during practice, but when access was removed, the GPT Base group scored 17% lower than students who never had access. These negative effects were largely mitigated by the safeguards in GPT Tutor.? The study suggests that the tool's design, and whether it lets users skip effortful thinking, matters more than whether AI is present.

3.3 Cognitive engagement (preliminary evidence)

A preprint from the MIT Media Lab assigned 54 adults to write essays with an LLM, with a search engine, or with no tools; 18 completed a fourth session in which conditions were swapped. Using electroencephalography (EEG), the authors reported the weakest connectivity in the LLM group, lower self-reported ownership of essays, and difficulty quoting one's own work.? The study is small, has not been peer reviewed, and EEG connectivity is not a direct measure of learning, so it should be treated as a hypothesis-generating signal rather than proof of harm.?

3.4 Depth of understanding

Melumad and Yun reported seven experiments with more than 10,000 participants in which people learned about topics either from LLM syntheses or from standard web search. LLM users reported shallower knowledge and wrote less informative advice, even when the underlying facts were identical, and the pattern persisted when web links were added to the LLM answers.? The authors' interpretation is that pre-synthesised answers turn learning from an active into a more passive activity. They do not conclude that LLMs are harmful in all uses; effects depend on how the tool is used.?

3.5 Using established learning science with AI

One of the best-replicated findings in memory research is the testing effect. In Roediger and Karpicke's experiments, restudying helped most on a test five minutes later, but on tests two days and one week later, students who had taken a recall test retained more than those who restudied, even though restudying made them feel more confident.[12] An application, consistent with this evidence, is to use AI to *quiz* you rather than to *summarise for* you.

Practical implication (author's synthesis): attempt the problem or answer from memory first; use AI to explain what you got wrong; ask for questions rather than answers.

4. Researching Faster Without Fabricating

4.1 Fabricated and erroneous references

Walters and Wilder asked ChatGPT to write short literature reviews on 42 multidisciplinary topics and checked all 636 citations in the 84 resulting papers. Of GPT-3.5 citations, 55% were fabricated, compared with 18% for GPT-4. Among the real (non-fabricated) citations, 43% of GPT-3.5 and 24% of GPT-4 citations contained substantive errors.? Newer tools may perform differently, but the safe assumption for beginners is that any reference produced from a chatbot's memory is unverified until found in a real database.

4.2 Even published evidence needs checking

A 2025 meta-analysis of 51 studies reported a large positive effect of ChatGPT on learning performance (Hedges' g = 0.867).? The journal retracted it on 22 April 2026, citing discrepancies in the meta-analysis that undermined confidence in its validity and conclusions; the authors did not respond to correspondence about the retraction.[1]? The case shows why a single headline number, however widely shared, should not be treated as settled. It also does not show the opposite: its retraction removes one pooled estimate and does not demonstrate that AI use has no educational value.

4.3 Do not rely on AI detectors

Liang and colleagues tested seven AI-text detectors on 91 essays written by non-native English speakers for the TOEFL exam and 88 US eighth-grade essays. The detectors were near-accurate on the eighth-grade essays but misclassified more than half of the TOEFL essays as AI-generated; all seven flagged 19.78% of them (18 of 91), and 89 of 91 (97.80%) were flagged by at least one detector.? For students, this means that a detector score is weak evidence of anything, and that keeping drafts and notes is a better protection than trying to "pass" a detector.

5. A Beginner's Workflow: Draft–Check–Own

The workflow below combines the evidence above with practical judgement. Entries without superscripts are the author's recommendations rather than findings from a specific study.

Table 1. Draft–Check–Own workflow by task

Task	Good use of AI	Main risk	Verification step
Writing	Outline, rephrase, adjust tone, feedback on clarity[3]	Accepting output with little editing; weaker ownership?	Write your key argument first; edit line by line; check every fact

Task	Good use of AI	Main risk	Verification step
Studying	Explain a concept, generate practice questions	Passive learning and over-reliance?,?	Attempt answers from memory before asking;[12] compare with your textbook or lecture notes
Literature search	Brainstorm keywords and synonyms; explain terminology	Fabricated or erroneous references?	Find every paper in PubMed, Google Scholar or the publisher's site; confirm authors, year and DOI; read the abstract
Reading papers	Plain-language overview as a map	Misrepresented findings	Read methods and results yourself; check for retractions[1]?
Academic integrity	Follow your institution's policy; keep drafts and prompts	Detector false positives?	Do not treat detector scores as proof

Three starter prompts that match the evidence (the first two build in the kind of safeguard that distinguished GPT Tutor from GPT Base?):

• *Writing:* "Here is my draft paragraph. Point out unclear sentences and gaps in logic, but do not rewrite it."

• *Studying:* "Quiz me on this chapter one question at a time. Wait for my answer before giving feedback."

• *Research:* "Suggest search keywords and synonyms for this topic. Do not list papers or references."

Statistical Tables and Comparative Analysis

Table 2. Reported AI use among full-time UK undergraduates (HEPI/Kortext surveys)

Indicator	2024 edition	2025 edition	2026 edition
Use AI in some form	66%[1,2]	92%[2]	95%[1]
Use generative AI for assessed work	53%[2]	88%[2]	94%[1]
Sample size (full-time undergraduates)	Not reported here	1,041[2]	1,054[1]

Notes: self-reported, UK-only data; sample sizes and question wording may differ between editions, so year-to-year comparisons are indicative rather than exact. The 2024 values are as reported in the later editions.

Table 3. Summary of key studies discussed

Study	Design and sample	Main finding	Key limitation
Noy and Zhang, 2023[3]	Preregistered online randomised experiment; 453 professionals; GPT-3.5	Time ?40%; rated quality +18%	Short online tasks; older model
Bastani et al., 2025?	Field experiment; nearly 1,000 high school maths students; GPT-4 tutors	Practice performance improved; GPT Base group 17% lower after access removed; safeguards largely mitigated harm	One subject and setting
Kosmyna et al., 2025?	Preprint; 54 adults, three groups; EEG; 18 in fourth session	Weakest connectivity and poorest self-quoting in LLM group	Small; not peer reviewed; EEG is indirect
Melumad and Yun, 2025?	Seven experiments; more than 10,000 participants	Shallower knowledge from LLM summaries than web search	Short-term tasks; effect depends on use

Study	Design and sample	Main finding	Key limitation
Walters and Wilder, 2023?	84 AI-written reviews; 636 citations	Fabricated: 55% (GPT-3.5), 18% (GPT-4)	Older models
Liang et al., 2023?	Seven detectors; 91 TOEFL and 88 US eighth-grade essays	Over half of TOEFL essays misclassified as AI	Detectors from 2023
Wang and Fan, 2025?	Meta-analysis of 51 studies	Reported g = 0.867; **retracted April 2026**[1]?	Not valid evidence

Suggested figure (Figure 1). A grouped bar chart of the Table 2 percentages (any AI use and AI use for assessed work) for 2024, 2025 and 2026, captioned “Self-reported AI use among UK full-time undergraduates, HEPI/Kortext surveys,” with the sample sizes and the UK-only limitation stated in the caption.

Discussion and Critical Analysis

Established, with caveats. Controlled evidence supports large time savings and quality gains for short professional writing tasks,[3] and audits show that chatbots in the tested versions fabricated or distorted many references.? These two findings are reasonably consistent with each other: AI is fast at producing text and unreliable as a source of verified facts.

Emerging. The learning evidence points to a distinction between *performance while using the tool* and *learning retained afterwards*. Bastani et al. show this directly,? Melumad and Yun show shallower knowledge after LLM-synthesised answers,? and the MIT preprint is directionally similar but preliminary.? The strength of the inference differs: the first two are peer-reviewed experiments, whereas the third is a small unreviewed study. None of them shows that AI use *necessarily* harms learning, and the GPT Tutor result suggests that design and instructions can reduce the risk.?

Contested or uncertain. The net effect of AI on learning across populations remains unsettled. The retraction of a prominent meta-analysis[1]? means that a pooled “large positive effect” should not be assumed. Other syntheses exist but were not assessed in this review, and the primary studies are mostly short-term, use earlier models, and vary widely in design.

Correlation versus causation. Survey data[1,2] show how many students use AI, not whether use improves grades. Experiments[3],?,? support causal claims within their tasks, but their findings may not generalise to different subjects, ages, or newer tools.

Equity. Detector bias against non-native English writers? is a concrete example of how AI-related policies can disadvantage some students. Unequal access to training is another: only 36% of surveyed UK students reported institutional AI-skills training.[2]

Future Perspectives and Research Directions

Several questions remain open. Longer-term, preregistered studies are needed to test whether guardrailed AI tutors improve retention over months rather than days. The preliminary EEG findings need peer review and replication with larger samples and behavioural outcomes. Reference-fabrication audits should be repeated on current tools, including those that retrieve and cite sources. Evidence from India and other non-UK settings is needed, including multilingual writing and study use. Finally, assessment design and AI-literacy teaching are likely to matter as much as the tools themselves, a point echoed by the HEPI authors’ recommendation that institutions stress-test assessments.[2]

Conclusion

The evidence supports a cautious but practical position. For writing, AI tools can reduce time and raise rated quality on short tasks.[3] For studying, the benefits depend on whether the tool supports effortful thinking or replaces it.?,? For research, every AI-supplied reference or factual claim must be verified in an independent source.? Published claims, including headline meta-analyses, can also be wrong,?,[1]? and AI detectors cannot be

treated as reliable arbiters of authorship.? A beginner who drafts with AI, checks everything, and keeps ownership of the core thinking can capture much of the speed benefit while limiting the main risks. Because tools and evidence are changing quickly, the workflow should be revisited as new peer-reviewed studies appear.

Competing interests

The authors declare that they have no competing interests.

References

1. Stephenson R, Armstrong C. Student Generative AI Survey 2026. HEPI Report 199. Oxford: Higher Education Policy Institute; 2026 Mar. Available from: https://www.hepi.ac.uk/2026/03/12/new-hepi-and-kortext-report-student-generative-artificial-intelligence-survey-2026/ [cited 2026 Nov 26].
2. Freeman J. Student Generative AI Survey 2025. HEPI Policy Note 61. Oxford: Higher Education Policy Institute; 2025 Feb. Available from: https://www.hepi.ac.uk/wp-content/uploads/2025/02/HEPI-Policy-Note-61.pdf [cited 2026 Nov 26].
3. Noy S, Zhang W. Experimental evidence on the productivity effects of generative artificial intelligence. Science. 2023;381(6654):187-92. doi: 10.1126/science.adh2586.
4. Bastani H, Bastani O, Sungu A, Ge H, Kabakc? Ö, Mariman R. Generative AI without guardrails can harm learning: evidence from high school mathematics. Proc Natl Acad Sci U S A. 2025;122(26):e2422633122. doi: 10.1073/pnas.2422633122.
5. Kosmyna N, Hauptmann E, Yuan YT, Situ J, Liao XH, Beresnitzky AV, et al. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv:2506.08872 [preprint, not peer reviewed]. 2025. doi: 10.48550/arXiv.2506.08872.
6. Melumad S, Yun JH. Experimental evidence of the effects of large language models versus web search on depth of learning. PNAS Nexus. 2025;4(10):pgaf316. doi: 10.1093/pnasnexus/pgaf316.
7. Walters WH, Wilder EI. Fabrication and errors in the bibliographic citations generated by ChatGPT. Sci Rep. 2023;13:14045. doi: 10.1038/s41598-023-41032-5.

How to cite this article (Vancouver)
Jha A, Gupta M, Nakamura Y. How to Use AI Tools to Write, Study and Research Faster: An Evidence-Based Beginner's Guide. CHEARS Chronical. 2026;13(4):e240. Available from: https://chears.in/journal/volume-13/issue-4/how-to-use-ai-tools-to-write-study-and-research-faster-an-evidence-based-beginne-240
Published in CHEARS Chronical.
https://chears.in/journal/volume-13/issue-4/how-to-use-ai-tools-to-write-study-and-research-faster-an-evidence-based-beginne-240