Conceptual Growth or Cognitive Shortcut? Examining GenAI's Impact on Statistical Learning Among Life Science Undergraduates: An Annotated Bibliography

Victoria Navarrete

Department of Statistical Sciences, University of Toronto

STA496H1Y: Readings in Statistics

Professor Bethany White

August 22, 2026

INTRODUCTION

Student usage of generative AI (genAI) in quantitative disciplines has exploded since OpenAI's 2022 release of ChatGPT (Marr, 2023). As context for the research focus that follows, a leading application among undergraduates is mastering course content in, supposedly, a more effective manner compared to traditional methods of learning. The integration of genAI into undergraduates' learning habits, on one hand, allows individuals to incorporate unique specificities directly into previously one-size-fits-all learning materials. This can facilitate increased enthusiasm and active participation in the learning process, which some studies suggest translates into stronger engagement with the material and faster learning (Binhammad et al., 2024). Conversely, as attention towards the accelerated learning process provided by ChatGPT and other large language models (LLMs) has heightened, a growing body of peer-reviewed research flag patterns of dependency. For instance, genAI tools are now being recognised as enabling students to offload metacognitive effort, hindering deep learning and contributing to what researchers have coined "cognitive laziness" (Yunus et al., 2025). These competing outcomes matter most in domains where students already struggle, and statistics is a persistent case. There is a consistent framing of statistics as a discipline that life science students approach with low confidence and motivation, especially at the undergraduate level (Kulacki & Aikens, 2026), which is a pattern that recurs whether or not genAI is part of the discussion. Because statistics is already a challenging subject for many students, genAI has the potential to either strengthen or harm life science students' development of genuine conceptual understanding.

It is this tension that motivates the annotated bibliography's focus. The research topic of interest is the tension between genAI's potential for improving statistical learning outcomes and the risk of introducing cognitive laziness. Specifically, I am interested in the extent to which undergraduate life science students' usage of genAI (focusing on ChatGPT) as a personalised learning support for statistics fosters conceptual understanding versus cognitive laziness, and how the way genAI is incorporated into coursework affects which outcome predominates.

This focus serves the purpose of determining whether genAI is closing or widening the gap among life science undergraduates. If left unaddressed, this gap can persist beyond the undergraduate learning phase into research practice itself: research output errors are often linked to inappropriate applications of statistics, a problem that persists even in high-impact factor publications (White & Singh, 2023). Therefore, better understanding genAI's effect on learning outcomes serves a purpose on two fronts: it informs the statistical education community on whether genAI has the potential to improve statistical literacy and thereby reduce research output errors, and, if so, how it should be incorporated into coursework to achieve this while avoiding the risk of cognitive laziness. More broadly, this research focus carries real-world significance: life science research directly informs medical and clinical conclusions. As such, understanding whether genAI is strengthening or further eroding the undergraduate statistical literacy gap is not merely significant as an educational goal, but as a safeguard against flawed evidence stemming from statistical errors shaping real-world health decisions.

ANNOTATED BIBLIOGRAPHY

Ahmed, M., Zaid, N. M., & Abdullah, A. H. B. (2025). The impact of integration of AI in statistics learning on student motivation and learning engagement in higher education sector. ***International Journal of Academic Research in Progressive Education and Development*****,** ***14*****(1), 426–436**

This quantitative study explored how student motivation and learning engagement in undergraduate statistics courses is impacted by ChatGPT, by conducting a regression analysis on survey data from 121 Abu Dhabi University students. The results indicate ChatGPT has a statistically significant positive effect on both motivation ($\beta = 0.329$, $p < .001$) and learning engagement ($\beta = 0.372$, $p < .001$), and authors link these positive learning effects to Self-Determination Theory: ChatGPT provides autonomy-supportive feedback which satisfies student psychological needs for competence and relatedness. The source is peer-reviewed and clearly defines its research methodology, suggesting credibility. Despite this, its metrics for motivation and engagement are self reported via survey, making it difficult to objectively assess conceptual statistical understanding. The source additionally provides empirical, quantitative evidence to support the personalisation-driven benefits side of the genAI tension I am researching: based on this source, individualised AI support may foster statistically significant positive engagement with statistics compared to non-ChatGPT using undergraduates. However, its exclusive focus on motivational outcomes, without addressing whether this engagement translates into genuine conceptual grasp compared to the risk of creating underlying dependence, raises an important gap other annotated bibliography sources must further investigate.

Alshboul, H. (2026). AI cognitive laziness in programming learning among university students. ***Journal of Humanities and Social Sciences Studies, 8*****(4), 26-30.**

This quantitative paper surveys 105 undergraduate programming students to measure how reliance on genAI tools connects with cognitive laziness. Moreover, programming is not merely parallel to statistics but often embedded within statistical concepts itself (for instance, regression analysis and hypothesis testing); the genAI-impacted programming learning this study examines is relevant to statistical conceptual understanding. Leveraging a 20-item Likert-scale survey and an AI Dependency Index, results indicated heightened cognitive laziness scores ($M = 3.42$) and heightened genAI dependency scores (3.65/5). The source moreover showed diminished independent problem-solving effort, even as students reported the tools supported their grasp of underlying concepts, a split directly relevant to the research topic's first portion: conceptual-understanding-versus cognitive-laziness-tension. Alshboul's call for scaffold-based integration of genAI tools into programming education over substitution of traditional learning methods informs research question's second portion: how the method of genAI coursework

incorporation determines which outcome (conceptual growth or cognitive laziness) predominates. As a peer-reviewed study investigating programming (which is deeply intertwined with statistics), this source presents an empirical lens on genAI-impacted statistics learning that much existing literature discusses only qualitatively, strengthening overall research.

Asy'ari, M., & Sharov, S. (2024). Transforming education with ChatGPT: Advancing personalized learning, accessibility, and ethical AI integration. ***International Journal of Essential Competencies in Education, 3*****(2), 119-157**

This systematic review uses a second-hand research approach: a literature review of over 1,000 Scopus-indexed publications, to identify findings on ChatGPT's role in education. Personalised learning is identified as a key theme: the review finds ChatGPT generates tailored practice materials, individualised feedback, and adjusts content complexity based on individual learning pace. Moreover, the source notes that the tool's feature of individualised responsiveness is linked with improved comprehension and conceptual retention. The review acknowledges recurring concerns of over-reliance and weakened critical thinking throughout the same body of literature, strengthening the research focus's exact tension between genAI supporting personalised learning and worsening cognitive laziness. While the literature review investigates genAI's impact on education generally, one publication within its set, Xing (2024, as cited in Asy'ari & Sharov, 2024), specifically looks at ChatGPT's applications in statistics instruction, which provides further relevance to the annotated bibliography's overall emphasis on statistics education impacts. Combining findings from several publications, this source presents high-level qualitative findings distinct from other empirical sources present in this annotated bibliography. Crucial to the research focus, the review reveals that genAI's personalization-versus-laziness tension persists across disciplines, validating interdisciplinary research relevance for the impact of genAI on statistics learning within the annotated bibliography.

Awunkori, E., & Matta, V. (2026). Statistics education rebooted in the age of artificial intelligence. ***Journal of Decision Systems, 35*****(1), 2657509.**

This systematic review by Awunkori and Matta (2026) synthesises 20 peer-reviewed studies (2020-2026) exploring the integration of AI, including genAI, into statistics education. The authors examine two relevant research questions: how AI-enabled tools are being integrated into statistics teaching and how AI-supported instruction affects conceptual understanding and statistical reasoning. The review suggests AI-enabled tools serve three recurring uses: personalised learning, on-demand conceptual clarification, and problem-solving guidance. Crucially, the study finds genAI can contribute to increased statistical comprehension, but these benefits are conditional on incorporation method: when genAI is used to bypass analytical engagement, it instead weakens statistical reasoning.

This paper directly informs my focus on genAI's tension between the potential to improve learning outcomes for undergraduate life science students and its cognitive laziness risk. Rather than examining a single tool or context, it surveys the landscape of instructional uses and learning outcomes across 20 studies, finding that genAI strengthens conceptual grasp only when coursework requires active, as opposed to passive, engagement with AI output. This provides a foundation for what is already known regarding genAI integration into coursework and how it affects learning outcomes, before I narrow toward more specific questions about conceptual grasp, ChatGPT, and the life sciences context.

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, 56*****(2), 489–530. https://doi.org/10.1111/bjet.13544**

This randomised experimental study by Fan et al. (2025) tracks 117 undergraduate students in a lab setting to examine the impact of genAI usage in a writing task, by testing how groups performed with different tools (genAI, human expert, or checklist). The source investigates how genAI affects three key learning areas: self-regulated learning (SRL) processes, learning performances, and learner's motivations. While intrinsic motivation did not differ with genAI usage, the study found SRL processes centered around genAI tools themselves, usage may contribute to undergraduate dependence and introduce "cognitive laziness" risk. While the genAI-using group significantly outperformed human expert and checklist groups in writing task performance, there was no difference in knowledge gain, which the authors argue is evidence of genAI's short-term task performance boost without learning improvements. This directly informs the cognitive laziness half of my research tension. The study's use of behavioural tracking strengthens its contribution, particularly because many other sources in this annotated bibliography rely on inherently subjective self-reported measures. The paper provides objective evidence that genAI use can improve measured performance while simultaneously promoting harmful cognitive offloading.

Febriyanti, F., Muntoha, T., & Uminar, A. N. (2026). The Ghost Learner Paradox: A Systematic Review of Metacognitive Laziness in the Age of Generative AI. ***Journal of Socio-Digital Integration, 1*****(1), 1-13.**

This PRISMA-guided systematic review synthesises peer-reviewed literature (2023-2026) to investigate the "Ghost Learner" paradox, where student performance in performance-based "learning" metrics improve while true cognitive development declines. The authors consolidate

Fan et al. (2025) and others to find that beneficial offloading (freeing working memory for higher-order reasoning) and detrimental substitution (bypassing productive struggle and contributing to cognitive laziness) are critically distinct outcomes resulting from genAI tool usage. The review suggests that current output-centric learning assessments fail to distinguish between beneficial offloading or detrimental substitution, leading to learning gaps between students using genAI tools. Additionally, genAI may widen Dunning-Kruger gaps between student's perceived ability even when actual conceptual grasp lags.

This paper is relevant to my research topic of genAI's potential for strengthening undergraduate life science learning outcomes compared to cognitive laziness risk, while also proposing genAI integration methods. Its distinction between beneficial offloading, which frees cognitive bandwidth for higher-order statistical learning, and detrimental substitution, which reduces cognitive engagement without being detected by conventional statistics assessments, is particularly relevant. Overall, the review serves as a mechanism for why apparent statistical learning gains may be overstated when genAI substitutes for student's own reasoning, and proposes process-oriented coursework to preserve cognitive engagement.

Fidalgo-Blanco, A., Fonseca-Escudero, D., & Sein-Echaluce, M. L. (2026). An integrated model based on generative AI for personalized learning and formative feedback. ***RIED-Revista Iberoamericana de Educación a Distancia, 29*****(2), 209-233**

This mixed-methods study by Fidalgo-Blanco et al. (2026) implements and evaluates the "Locate, Transform, Consolidate" model by embedding genAI into a first-year undergraduate mathematics course notorious for high failure rates, focusing on its potential as a personalisable learning support. The authors compared a control and experimental group with the same baseline motivation, finding that the experimental group scored 1.04 SD higher on a surprise test; students attributing these gains to genAI's personalisation of learning materials and individual-specific feedback, the same features that distinguish genAI as a personalised learning support rather than a cognitive substitute.

This paper strongly supports the "improved learning outcomes" half of my research tension: personalised support, incorporated into coursework as a pedagogical mediator rather than an answer-generator, correlates with deeper student engagement in a challenging first-year quantitative subject analogous to introductory statistics for life science undergraduates. Importantly, usage "did not stand out as a separate positive or negative pole", suggesting genAI is capable of being successfully integrated into student learning environments. This study directly guides the question about the coursework-implementation component of my research topic by providing an evidence-based contrast between unsupervised ChatGPT use currently undergraduate statistics courses, where cognitive laziness may be more present.

Khosravi, H., Shibani, A., Jovanovic, J., A. Pardos, Z., & Yan, L. (2025). Generative AI and Learning Analytics: Pushing Boundaries, Preserving Principles. ***Journal of Learning Analytics, 12*****(1), 1–11. https://doi.org/10.18608/jla.2025.8961**

Khosravi et al. (2025) provide a conceptual synthesis of 10 papers on genAI and learning analytics, organised under the framework of Clow's learning-analytics (LA) cycle. The authors focused on four key areas: understanding learning contexts, harnessing genAI for learning improvements, applying LA methods to create meaningful learning insights, and designing interventions that optimise genAI learning outcomes. The authors suggest that genAI allows surface-level task performance while offloading the cognitive effort required for lasting conceptual learning, potentially producing "cognitive laziness". Relevant to my research focus, the study summarises Lai et al.'s study of Socratic chatbot usage by 34 students in an introductory statistics course, through which epistemic network analysis distinguished reflective genAI engagement from answering-seeking.

Notably for my research focus, this suggests that identical genAI learning supports lead to different learning outcomes depending on how students engage with it. The study therefore emphasises genAI learning analytics that support captures students' reasoning processes, rather than final performance alone: incorporating coursework that promotes cognitive effort and self-regulation may promote this. This directly informs the research topic of whether ChatGPT in statistics serves as personalized scaffolding or cognitive offloading, and how intentional coursework design can shift students toward the former.

Pardos, Z. A., & Bhandari, S. (2024). ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. ***PLoS One, 19*****(5), e0304013**

Pardos and Bhandari (2024) provide unusually direct evidence that genAI can foster conceptual statistical understanding through personalised learning support, rather than hinder it. Using a randomised study of 274 learners, ChatGPT was compared to human tutors across randomly assigned subject areas (statistics, college algebra, intermediate algebra and elementary algebra). ChatGPT-generated hints resulted in a 17% overall learning gain, exceeding the no-hint control and not differing significantly in human tutor-based learning gains. Relevant for my research focus, the source independently assessed statistics: ChatGPT-supported learning improved by 18.83% ($p = 0.007$), compared to 11.10% using human tutors and -7.56% without help. The study's findings provide a relevant counterbalance to my research topic's cognitive laziness risk: when genAI is scaffolded as a supplementary learning aid within coursework, it creates learning gain equivalent to human assistance, including in statistics. While the study doesn't investigate personalized dialogue features, authors assert genAI could provide more adaptive and time-efficient tutoring, since many LLMS can adapt based on students' previous prompts and

existing conceptual gaps. The comparable effectiveness of human and genAI tutoring suggests personalised statistical support could be extended beyond students who can afford private tutoring, showing potential as an accessible personalised learning support.

Salih, S., Husain, O., Hamdan, M., Abdelsalam, S., Elshafie, H., & Motwakel, A. (2025). Transforming education with AI: A systematic review of ChatGPT's role in learning, academic practices, and References - 9 institutional adoption. *Results in Engineering, 25,* 103837. https://doi.org/10.1016/j.rineng.2024.103837

This systematic review of AI-supported teaching methods provides a strong synthesis for the cognitive-laziness side of my research question. The study uses PRISMA procedures to search Science Direct, IEEE, Web of Science, Scopus, and Springer for studies published between 2022 and 2024; two authors independently screened studies, resolving discrepancies via discussion and third-reviewer consulting. Synthesising 40 selected papers, the source identified over-reliance as a significant educational risk: because genAI tool usage often allows students to outsource reflection, autonomous reasoning, and problem-solving processes, this risks weakening students' critical-thinking skills. Within this source, this conclusion is further evidenced by inclusion of studies including two 2024 studies (a systematic review and qualitative study) which used analysis of AI dialogue systems and post-test assessment respectively to analyse genAI's impact on students' cognitive abilities, both associating over-reliance of these tools with reduced critical thinking and problem-solving and asserting the importance of balanced genAI strategies in educational institutions. Notably with regards to my research focus, Salih et al. emphasise that genAI's effects on cognition are mitigable: design and implementation of policies that encourage active learning and human supervision as opposed to answer-seeking may allow for genAI to serve as a personalisable tool without jeopardizing cognitive development.

Wahba, F., Ajlouni, A. O., & Abumosa, M. A. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(7), em2468

Wahba et al. use a quasi-experimental pretest–posttest design to test whether ChatGPT-based statistics instruction improves undergraduates' statistical reasoning and attitudes. Fifty-six BSc elementary-education students were enrolled in two existing course sections, which were randomly assigned to experimental or control conditions (n=28 each). Across three weeks, both groups received the same instructor, content, Moodle materials, four homework tasks, and four collaborative activities; the experimental group additionally received ChatGPT training and used

it for statistical analysis, hypothesis testing, probability, and problem-solving. Outcomes were measured using Garfield's validated 20-item Statistical Reasoning Assessment (SRA; α=.82) and the validated 28-item Survey of Attitudes Toward Statistics (SATS-28; α=.79), both translated into Arabic, reviewed by eight specialists, and piloted with 45 students. ANCOVA showed significantly higher adjusted statistical-reasoning scores for ChatGPT users (11.24 vs. 9.25; F=11.85, p=.001, ηp^2=.183) and substantially more positive attitudes (F=372.78, p<.001, ηp^2=.876). This supports my focus on whether personalised ChatGPT assistance can foster conceptual statistical understanding, while its structured integration into otherwise equivalent coursework informs how implementation may shape learning outcomes. Its small, single-institution, short-duration sample limits generalisation to life-science undergraduates.

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*****(1), 28. https://doi.org/10.1186/s40561-024-00316-7**

This systematic review by Zhai et. al (2024) synthesises evidence connecting student over-reliance on AI dialogue system learning aids to reduced cognitive abilities, within educational and research settings. The authors examined four databases: ProQuest, IEEE Xplore, Web of Science and ScienceDirect for studies published from 2017 to 2023, ultimately analysing 14 peer-reviewed publications. The review suggests genAI tools impact student cognitive abilities in three pervasive ways: decision-making, critical-thinking, and analytical thinking. Relevant to the research focus, the review identifies genAI may establish a mechanism of cognitive laziness, where AI-generated answers provide time-efficient "shortcuts" that reduce the student incentive to independently work through cognitive processes. Notably, one case study of 245 students across 25 Indonesian universities reported concerns about impact on critical thinking (75%) and genAI reliance (73%). Authors assert that undergraduate coursework which encourages students to engage critically with AI output may offset this cognitive skill risk, which directly supports my research focus by proposing that genAI cognitive laziness risk is associated with poor genAI usage within coursework rather than tool access itself. Therefore, the study argues that genAI scaffolding into instructional design is critical to whether genAI more strongly serves as a personalised learning support or fosters cognitive laziness.

SYNTHESIS & CONCLUSION

Across the 12 annotated bibliography sources, the clearest overarching pattern with respect to findings is that genAI does not produce one binary educational outcome. Rather, genAI tools' impact appears strongly influenced by the mechanism with which students use it, and how instructors structure that use. Studies which emphasised the integration of personalised feedback, guided problem solving, pacing, and repeated examples when using genAI learning tools report stronger engagement, statistical reasoning, or measurable learning gains (Ahmed et al., 2025; Asy'ari & Sharov, 2024; Fidalgo-Blanco et al., 2026; Pardos & Bhandari, 2024; Wahba et al., 2024). By contrast, studies which centered on over-reliance report that genAI can significantly diminish independent problem-solving, critical thinking, metacognitive effort, and self-regulation (Alshboul, 2026; Fan et al., 2025; Salih et al., 2025; Zhai et al., 2024). As such, the insightful synthesis throughout sources is not that literature is divided between "pro-genAI" as a personalised learning support and "anti-genAI" given the risk of cognitive laziness; that overlooks the nuance related to how genAI tools are incorporated into coursework. Rather, both groups of studies point toward the same distinction: genAI is most beneficial when it serves as a learning tool to scaffold students' pre-existing cognitive processes and most harmful when it substitutes for it. This conditional pattern also is discussed explicitly in the broader reviews by Awunkori and Matta (2026) and Asy'ari and Sharov (2024).

In terms of the positive-learning literature which focuses on genAI's potential benefits as a personalisable tool, a noticeable pattern across many sources is that not all reported benefits constitute comparable evidence of boosted conceptual understanding. For instance, Ahmed et al. (2025) finds that genAI leads to higher motivation and student engagement: while this may support the learning process, this by itself does not demonstrate superior conceptual grasp of statistics. Providing more robust evidence, stronger performance on a validated Statistical Reasoning Assessment was seen with students using ChatGPT as an individualised learning tool compared with control students from Wabhba et al. (2024). Pardos and Bhandari (2024) also reinforce this pattern, finding that learning support from ChatGPT led to substantial mathematical learning gains, including for statistics, as well as Fidalgo-Blanco et al. (2026) which asserted a 1.04 SD improvement on a surprise mathematics assessment for those using genAI as learning tools. Collectively, the positive-leaning literature suggests that genAI tools are most beneficial to statistical learning outcomes when students use them alongside their own cognitive efforts (by prompting clarifications, asking to slow or speed up pace, or focus on more difficult concepts), as opposed to seeking instantaneous task completion.

Looking at the cognitive-laziness literature reveals interesting findings across sources as well. Here, sources agree that boosted learning performance metrics by genAI usage do not necessarily reflect improved conceptual grasp. The study by Fan et al. (2025) provides a relevant example: while student groups using genAI significantly outperformed those not using genAI, there was no measured difference in knowledge gain. This is potentially evidence of how genAI

can improve external performance such as essay writing while failing to create an equivalent improvement in conceptual understanding. Febriyanti et al. (2026) agrees on a similar problem, reporting the "Ghost Learner" paradox where outwards student performance in traditional "learning" metrics increase while internal cognitive development declines due to student over-reliance and cognitive offloading. Furthermore, Khosravi et al. (2025) argues that final outputs produced by students when using genAI tools may not reflect whether a student reflected through a task, and therefore may not accurately reflect how much cognitive gains (or lack thereof) was achieved by these tasks. Zhai et al. (2024), Salih et al. (2025), and Alshboul (2026) arrive at similar conclusions, despite different methodologies: over-reliance of genAI leads to pervasive declines in critical thinking, analytical abilities, and problem-solving. Prior to synthesising cognitive-laziness literature's findings, there was a gap between studies indicating potential gains in statistical learning and losses due to cognitive laziness. However, synthesizing cognitive-laziness literature's findings helps reconcile the apparent disagreement: these study groups are often measuring different outcomes stemming from different forms of genAI usage. As noted above, a student can become faster, more self-assured, or higher-scoring while deepening conceptual learning gaps.

Another notable pattern is what the literature is focusing on over time. Sources selected for the annotated bibliography span a two-year time range, from 2024 to 2026. Earlier works in 2024 tended to examine the role of genAI itself: Pardos and Bhandari (2024) investigate if ChatGPT could serve the role of a human tutor substitute, while Asy'ari and Sharov (2024) question ChatGPT's role in impacting education. By 2025 and 2026, sources increasingly shift their focus from the tool itself towards the mechanisms of use: self-regulation learning, cognitive offloading, scaffolding inside of curriculum, and instructor-supervised implementation (Fan et al., 2025; Khosravi et al., 2025; Awunkori & Matta, 2026; Fidalgo-Blanco et al., 2026). Therefore, a key theme seems to be literature moving from whether genAI plays a role in learning, to acknowledging a role and instead asking under what conditions genAI supports effective learning. This is a relevant shift in literature; it suggests an acknowledgement that genAI tools have a substantial impact on education acquired over time, and progress towards now positioning coursework design as the factor connecting tools to either conceptual growth or cognitive laziness. Approaching 2026, otherwise studies tended towards similar recommendations of human supervision, self-regulation, and coursework that encourages students to use their own cognitive processes alongside genAI tools.

References

Ahmed, M., Zaid, N. M., & Abdullah, A. H. B. (2025). The impact of integration of AI in statistics learning on student motivation and learning engagement in higher education sector. *International Journal of Academic Research in Progressive Education and Development*, *14*(1), 426–436

Alshboul, H. (2026). AI cognitive laziness in programming learning among university students. *Journal of Humanities and Social Sciences Studies, 8*(4), 26-30.

Asy'ari, M., & Sharov, S. (2024). Transforming education with ChatGPT: Advancing personalized learning, accessibility, and ethical AI integration. *International Journal of Essential Competencies in Education, 3*(2), 119-157

Awunkori, E., & Matta, V. (2026). Statistics education rebooted in the age of artificial intelligence. *Journal of Decision Systems, 35*(1), 2657509.

Binhammad, M. H. Y., Othman, A., Abuljadayel, L., Al Mheiri, H., Alkaabi, M., & Almarri, M. (2024). Investigating how generative AI can create personalized learning materials tailored to individual student needs. *Creative Education, 15*(7), 1499–1523. https://doi.org/10.4236/ce.2024.157091

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, 56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Febriyanti, F., Muntoha, T., & Uminar, A. N. (2026). The Ghost Learner Paradox: A Systematic Review of Metacognitive Laziness in the Age of Generative AI. *Journal of Socio-Digital Integration, 1*(1), 1-13.

Fidalgo-Blanco, A., Fonseca-Escudero, D., & Sein-Echaluce, M. L. (2026). An integrated model based on generative AI for personalized learning and formative feedback. *RIED-Revista Iberoamericana de Educación a Distancia, 29*(2), 209-233

Khosravi, H., Shibani, A., Jovanovic, J., A. Pardos, Z., & Yan, L. (2025). Generative AI and Learning Analytics: Pushing Boundaries, Preserving Principles. *Journal of Learning Analytics, 12*(1), 1–11. https://doi.org/10.18608/jla.2025.8961

Kulacki, A. R., & Aikens, M. L. (2026). Examining motivational attitudes toward statistics and their relationship to performance in life science students. *Journal of Statistics and Data Science Education, 34*(1), 36–47. https://doi.org/10.1080/26939169.2024.2365892

Marr, B. (2023, May 19). A short history of ChatGPT: How we got to where we are today. *Forbes.* https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/

Pardos, Z. A., & Bhandari, S. (2024). ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. *PLoS One, 19*(5), e0304013

Salih, S., Husain, O., Hamdan, M., Abdelsalam, S., Elshafie, H., & Motwakel, A. (2025). Transforming education with AI: A systematic review of ChatGPT's role in learning, academic practices, and References - 9 institutional adoption. *Results in Engineering, 25,* 103837. https://doi.org/10.1016/j.rineng.2024.103837

Wahba, F., Ajlouni, A. O., & Abumosa, M. A. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(7), em2468

White, B., & Singh, J. (2024). Preparing future life science researchers to engage with statistics in research. *IASE 2023 Satellite Conference – Fostering Learning of Statistics and Data Science.* https://doi.org/10.52041/iase2023.101

Yunus, A., Gay, P. R., & Lee, O. T. (2025). *From co-design to metacognitive laziness: Evaluating generative AI in vocational education* [Preprint]. arXiv. https://arxiv.org/abs/2512.12306

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*(1), 28. https://doi.org/10.1186/s40561-024-00316-7