American Journal of Applied Science and Technology

VOLUME Vol.06 Issue01 2026
PAGE NO. 153-160

Generative Artificial Intelligence, Critical Cognition, And Automated Test Engineering In Contemporary Education: A Multidimensional Theoretical And Empirical Synthesis

Julian T. Hargrave
Professor, University of Zurich, Switzerland

Received: 23 December 2025; **Accepted:** 12 January 2026; **Published:** 31 January 2026

Abstract: Generative artificial intelligence has become one of the most transformative technological developments affecting education, software engineering, and cognitive practice in the twenty first century. While large language models and generative systems are increasingly adopted for content creation, tutoring, assessment, and software development, their deeper implications for critical thinking, metacognition, professional competence, and epistemic trust remain under active debate. This study develops a comprehensive, theory driven and empirically grounded synthesis of how generative artificial intelligence, with a specific focus on automated behavior driven development and test engineering, interacts with human cognition and educational practice. Anchored in recent scholarship, the article integrates the emerging field of generative AI assisted software testing with classical and contemporary theories of learning, critical thinking, and human computer collaboration. Particular attention is given to the automation of behavior driven development through generative models, which has been shown to restructure how software specifications, test cases, and validation workflows are produced and interpreted, thereby altering the cognitive and organizational processes of engineering teams (Tiwari, 2025).

This research addresses a major gap in the current literature, which has tended to treat generative AI either as a productivity tool in engineering or as a pedagogical technology in education, but rarely as a socio technical system that simultaneously reshapes epistemic practices, critical judgment, and professional agency. By synthesizing insights from educational psychology, human computer interaction, ethical AI, and software engineering, this article constructs a unified framework for understanding generative AI not as a replacement for human reasoning but as a mixed initiative partner that co constructs meaning, standards, and decisions. The methodological approach is interpretive and integrative, drawing on structured qualitative synthesis of the provided references and applying them to the specific domain of automated testing and behavior driven development. This allows for a rigorous exploration of how generative AI systems mediate cognition, influence trust, and modify institutional norms in both classrooms and development environments.

The findings indicate that while generative AI significantly enhances efficiency, consistency, and coverage in test automation and instructional design, it also introduces risks of cognitive offloading, metacognitive erosion, and epistemic overreliance. Studies of critical thinking and human AI collaboration suggest that performance gains do not automatically translate into deeper understanding or reflective awareness, a phenomenon that becomes particularly visible when engineers or students rely on generative systems to produce complex artifacts without fully engaging in the underlying reasoning (Fernandes et al., 2024; Facione et al., 2011). The article argues that the future of generative AI in education and software engineering depends on the development of pedagogical and organizational scaffolds that maintain human agency, promote reflective interaction, and align technological

automation with the goals of critical inquiry and responsible innovation. By embedding the automation of behavior driven development within a broader theory of human AI cognition, this study provides a foundation for future research, policy, and practice in the evolving landscape of generative artificial intelligence.

Keywords: Generative artificial intelligence, behavior driven development, critical thinking, educational technology, human AI collaboration, test automation.

Introduction: The rapid diffusion of generative artificial intelligence across educational and professional domains has initiated a profound transformation in how knowledge is produced, validated, and applied. Unlike earlier forms of educational technology or software automation, generative AI systems such as large language models operate not merely as tools for retrieval or calculation but as semi autonomous agents capable of producing natural language, code, and structured artifacts that resemble human outputs in both form and complexity (Brown et al., 2020). This capability has led to their widespread adoption in domains as diverse as writing, tutoring, programming, and software testing, with significant implications for learning, labor, and epistemic authority (Ghosh et al., 2023; Celik et al., 2022). At the same time, scholars and practitioners have raised concerns about the cognitive, ethical, and educational consequences of delegating increasingly sophisticated tasks to generative systems, particularly when these systems mediate core activities such as reasoning, evaluation, and decision making (Bender et al., 2021; Adams et al., 2023).

One of the most salient but under theorized applications of generative AI lies in the automation of behavior driven development and test engineering. Behavior driven development is a collaborative software development methodology that emphasizes the use of natural language specifications to define system behavior, thereby aligning technical implementation with stakeholder understanding. Traditionally, this approach has relied on human authored scenarios, examples, and acceptance tests to bridge the gap between business requirements and executable code. However, recent advances in generative AI have made it possible to automatically generate, refine, and maintain these artifacts at scale, significantly increasing efficiency and consistency across the software lifecycle (Tiwari, 2025). This shift has important implications not only for software quality and productivity but also for how developers think about requirements, verification, and accountability.

The relevance of this development extends far beyond software engineering. As generative AI becomes integrated into educational platforms, assessment systems, and collaborative tools, similar dynamics of automation, cognitive offloading, and epistemic mediation are emerging in classrooms and training environments (Abunaseer, n.d.; Chauncey and McKenna, 2023). Students increasingly rely on AI to generate essays, solve problems, and summarize texts, while teachers use AI to design curricula, grade assignments, and provide feedback. These practices raise fundamental questions about the nature of learning, the cultivation of critical thinking, and the role of human judgment in an era of algorithmic assistance (Ahmad et al., 2023; Akgun and Greenhow, 2022).

Critical thinking, long regarded as a central goal of education and professional competence, provides a particularly important lens through which to analyze these changes. Theoretical frameworks developed by Facione and colleagues emphasize that critical thinking involves not only cognitive skills such as analysis, evaluation, and inference but also dispositions such as open mindedness, intellectual humility, and a willingness to engage with complexity (Facione, 1990; Facione et al., 2011). These qualities are cultivated through sustained engagement with problems, reflection on one's own reasoning, and dialogue with others. When generative AI systems intervene in these processes by providing ready made answers, explanations, or solutions, they may simultaneously enhance performance and undermine the metacognitive processes that support deep understanding, a tension that has been empirically documented in recent human AI interaction research (Fernandes et al., 2024; Göldi et al., 2024).

Within software engineering, the automation of behavior driven development exemplifies this tension in a particularly acute form. On the one hand, generative AI can dramatically reduce the time and effort required to produce comprehensive test suites, translate requirements into executable scenarios, and maintain alignment between code and specification, thereby addressing long standing challenges of quality assurance and technical debt (Tiwari, 2025). On the other hand, when engineers rely on AI generated tests and specifications, they may engage less deeply with the underlying business logic, edge cases, and ethical implications of their systems, potentially weakening their capacity for critical evaluation and responsible design (Glikson and Woolley, 2020; Gordon et al.,

2023).

Despite the growing importance of these issues, the existing literature remains fragmented. Studies of generative AI in education often focus on pedagogical outcomes, ethical concerns, or academic integrity without connecting these to the technical realities of AI assisted development and testing (Cotton et al., 2023; Clark et al., 2021). Conversely, research on AI driven software engineering tends to emphasize efficiency and performance while neglecting the cognitive and educational dimensions of human AI collaboration (Tiwari, 2025; Horvitz, 1999). There is therefore a pressing need for an integrative framework that situates the automation of behavior driven development within a broader theory of learning, cognition, and human computer interaction.

This article addresses that gap by developing a comprehensive synthesis of how generative artificial intelligence reshapes critical thinking, metacognition, and professional practice through its role in automated test engineering and educational technology. Drawing exclusively on the provided body of scholarly references, the study weaves together insights from educational psychology, ethical AI, human computer interaction, and software engineering to construct a multidimensional account of generative AI as a socio technical system. In doing so, it advances three interrelated arguments. First, that generative AI driven automation of behavior driven development represents not merely a technical innovation but a reconfiguration of epistemic practices in which human understanding and machine generation become deeply intertwined (Tiwari, 2025). Second, that the cognitive and educational impacts of this reconfiguration are best understood through the lens of critical thinking and metacognition, which reveal both the opportunities and risks of AI mediated performance (Facione et al., 1995; Fernandes et al., 2024). Third, that the ethical and pedagogical challenges posed by generative AI require not simple restriction or uncritical adoption but the design of mixed initiative systems and reflective scaffolds that preserve human agency and accountability (Chauncey and McKenna, 2023; Horvitz, 1999).

By situating automated behavior driven development within this broader theoretical landscape, the article contributes to a more holistic understanding of generative AI in contemporary education and professional practice. It provides a foundation for future empirical research, curriculum design, and policy making that recognizes the complex interplay between technological automation and human cognition in the age of generative artificial intelligence.

METHODOLOGY

The methodological orientation of this study is grounded in a qualitative, theory driven synthesis of the scholarly literature on generative artificial intelligence, education, critical thinking, and automated software testing. Given the rapid evolution of generative AI technologies and the ethical and cognitive complexities they introduce, a purely quantitative or experimental approach would be insufficient to capture the depth and interconnection of the phenomena under investigation (Celik et al., 2022; Clark et al., 2021). Instead, this research adopts an interpretive integrative methodology that treats the provided references as a coherent corpus of theoretical and empirical insights from which a comprehensive analytical framework can be constructed.

At the core of this methodology lies the systematic integration of perspectives from multiple disciplinary domains. Educational psychology contributes theories of learning, critical thinking, and metacognition, particularly through the work of Facione and colleagues, whose Delphi Report and subsequent studies establish a widely accepted conceptualization of critical thinking as a combination of cognitive skills and dispositional orientations (Facione, 1990; Facione et al., 2011). Human computer interaction and mixed initiative systems provide a second pillar, offering models of how humans and intelligent systems collaborate, negotiate control, and maintain trust in shared tasks (Horvitz, 1999; Glikson and Woolley, 2020). Software engineering and test automation, as articulated in recent work on generative AI assisted behavior driven development, supply the third pillar, grounding the analysis in concrete practices of specification, verification, and quality assurance (Tiwari, 2025).

The methodological process begins with a close reading of each reference to identify its core theoretical claims, empirical findings, and conceptual frameworks. Rather than treating these sources as discrete studies, the analysis seeks to map their points of convergence and divergence, particularly where they address common themes such as human agency, cognitive effort, ethical responsibility, and the role of automation. For example, studies of AI in education that highlight the risks of student laziness or diminished decision making are read alongside research on digital amnesia and overreliance on technology, revealing a broader pattern of cognitive offloading in technologically mediated environments (Ahmad et al., 2023; Greenwood and Quinn, 2017). Similarly, work on AI supported writing and debate scaffolding is juxtaposed with findings on metacognitive disconnects in human AI collaboration to illuminate how performance gains

can mask deeper deficits in reflective awareness (Holzer et al., 2018; Fernandes et al., 2024).

A key methodological innovation of this study is the explicit integration of automated behavior driven development into this cognitive and educational framework. Drawing on the detailed account of generative AI driven test automation provided by Tiwari (2025), the analysis treats software testing not merely as a technical activity but as a form of applied reasoning and learning in which developers interpret requirements, anticipate user behavior, and evaluate system responses. By conceptualizing behavior driven development as an epistemic practice, the methodology allows insights from critical thinking theory and educational research to be directly applied to the domain of software engineering.

The study employs a form of qualitative meta synthesis in which themes, constructs, and mechanisms are iteratively refined through comparison across sources. For instance, the notion of trust in AI systems, as discussed by Glikson and Woolley (2020), is examined in relation to ethical frameworks for AI use in education (Adams et al., 2023; Akgun and Greenhow, 2022) and to practical tools for auditing and double checking AI generated content (Gordon et al., 2023). This triangulation enables a more nuanced understanding of how trust, verification, and accountability operate across different contexts of generative AI use.

While this methodology offers rich theoretical insight, it also has limitations that must be acknowledged. Because the analysis relies exclusively on the provided references, it cannot incorporate the full breadth of empirical data or technical detail that might be available in the wider literature. Moreover, the interpretive nature of the synthesis means that conclusions are shaped by the conceptual lenses applied, particularly the emphasis on critical thinking and human AI collaboration. Nevertheless, within these constraints, the methodology provides a rigorous and coherent framework for examining the multifaceted impact of generative artificial intelligence on education and automated test engineering.

RESULTS

The integrative analysis of the provided literature yields a set of interrelated findings that illuminate how generative artificial intelligence reshapes cognition, professional practice, and educational outcomes through its role in automated behavior driven development and related applications. These findings are not presented as statistical outcomes but as theoretically grounded patterns that emerge from the convergence of multiple scholarly perspectives, reflecting the qualitative and interpretive nature of the methodology (Celik et al., 2022; Clark et al., 2021).

One of the most consistent results across the literature is that generative AI significantly enhances surface level performance and productivity in both educational and software engineering contexts. In automated behavior driven development, generative models can rapidly produce comprehensive test scenarios, translate natural language requirements into executable specifications, and adapt to changing codebases with a speed and consistency that far exceeds traditional manual methods (Tiwari, 2025). This efficiency gain mirrors findings in educational settings, where AI systems are able to generate explanations, summaries, and feedback that support students in completing assignments and mastering content more quickly (Abunaseer, n.d.; Brown et al., 2020). From an organizational perspective, these capabilities promise reductions in cost, time to market, and human error, reinforcing the appeal of generative AI as a transformative technology (Ghosh et al., 2023).

However, the literature also reveals a striking disconnect between performance improvements and deeper cognitive outcomes. Research on human AI collaboration demonstrates that users often achieve higher task performance when assisted by generative systems, yet show weaker metacognitive awareness and reduced ability to explain or justify their solutions independently (Fernandes et al., 2024; Göldi et al., 2024). This pattern is particularly relevant to automated testing, where developers may rely on AI generated test cases without fully understanding the underlying assumptions, coverage gaps, or potential biases embedded in those tests (Tiwari, 2025). In educational contexts, similar dynamics are observed when students use AI to produce essays or solve problems, leading to concerns about superficial learning and the erosion of critical thinking dispositions (Ahmad et al., 2023; Cotton et al., 2023).

A third major finding concerns the role of trust and epistemic authority in AI mediated environments. As generative systems produce increasingly plausible and authoritative outputs, users tend to attribute a high degree of credibility to their suggestions, sometimes even in the absence of transparent justification or verifiable evidence (Glikson and Woolley, 2020; Gordon et al., 2023). In behavior driven development, this can result in engineers accepting AI generated specifications or test results as correct without rigorous human validation, potentially allowing errors or misinterpretations to propagate through the development process (Tiwari, 2025). In education, similar patterns of overreliance raise ethical and pedagogical concerns, particularly when students or teachers treat AI outputs as definitive rather than

provisional and contestable (Chauncey and McKenna, 2023; Akgun and Greenhow, 2022).

The results also highlight the importance of mixed initiative interaction and scaffolding in mitigating these risks. Systems and pedagogical designs that encourage users to question, reflect on, and verify AI outputs are associated with stronger critical engagement and more responsible use of generative technologies (Horvitz, 1999; Holzer et al., 2018). For example, tools that provide contextual notifications about bias or uncertainty, or that prompt users to articulate their reasoning, can counteract the tendency toward passive acceptance and cognitive offloading (Harrison et al., 2024; Hauser and Schwarz, 2015). In the context of automated test engineering, this suggests that generative AI should be embedded within workflows that require human review, discussion, and iterative refinement of specifications and tests, rather than operating as an opaque, fully autonomous agent (Tiwari, 2025).

Collectively, these results indicate that the impact of generative artificial intelligence on education and software engineering is fundamentally ambivalent. While the technology delivers undeniable gains in efficiency and accessibility, it simultaneously challenges the cultivation of critical thinking, metacognition, and professional judgment that are essential for long term learning and responsible innovation (Facione et al., 1995; Fernandes et al., 2024). Understanding and addressing this tension is therefore central to the future development and governance of generative AI systems.

DISCUSSION

The findings of this study invite a deep theoretical and practical reflection on the role of generative artificial intelligence in shaping human cognition, professional practice, and educational values. At the heart of this reflection lies a fundamental paradox: generative AI systems, particularly when applied to automated behavior driven development and educational support, appear to make users more capable in terms of immediate task performance while simultaneously risking a weakening of the very cognitive and dispositional qualities that underpin genuine expertise and critical judgment (Fernandes et al., 2024; Facione et al., 2011). This paradox can be understood only by situating generative AI within a broader socio technical and epistemic framework that recognizes technology not as a neutral tool but as an active participant in the construction of knowledge and meaning.

From the perspective of software engineering, the automation of behavior driven development exemplifies how generative AI reshapes the epistemic foundations of technical work. Behavior driven development has traditionally been valued not only for its ability to produce executable specifications but also for its role in fostering shared understanding among developers, testers, and stakeholders through collaborative dialogue and example driven reasoning. By translating human language into structured tests, BDD creates a space in which assumptions can be surfaced, negotiated, and refined. When generative AI enters this space, as described by Tiwari (2025), it accelerates and scales the production of these artifacts, but it also mediates the dialogue itself, potentially narrowing the range of perspectives and interpretations that are brought to bear on a problem.

This mediation has important implications for critical thinking. According to the Delphi Report and subsequent work by Facione and colleagues, critical thinking is not merely the execution of cognitive operations but a reflective and self regulating process that involves questioning assumptions, evaluating evidence, and considering alternative explanations (Facione, 1990; Facione et al., 2011). In a generative AI driven BDD workflow, many of these operations are partially delegated to the system, which proposes scenarios, edge cases, and test structures based on patterns learned from vast corpora of code and text. While this can expose developers to possibilities they might not have considered, it can also reduce the need for them to actively generate and evaluate these possibilities themselves, thereby weakening the exercise of critical thinking skills over time.

The same tension is evident in educational contexts. Research on AI assisted learning shows that students often achieve higher grades or faster completion times when using generative tools, yet their ability to explain concepts, transfer knowledge to new situations, or reflect on their own understanding may decline (Ahmad et al., 2023; Fernandes et al., 2024). This phenomenon resonates with earlier concerns about digital amnesia and the outsourcing of memory and cognition to technological systems (Greenwood and Quinn, 2017). When generative AI provides ready made explanations and solutions, learners may bypass the struggle and uncertainty that are essential to deep learning, leading to what can be described as an epistemic hollowing out beneath the surface of high performance.

Trust plays a central role in this dynamic. As Glikson and Woolley (2020) have shown, humans tend to calibrate their trust in AI systems based on perceived competence and reliability, often attributing greater authority to systems that produce fluent and confident outputs. In the context of generative AI, this can lead to a form of epistemic deference in which users accept AI

generated content as correct or optimal without sufficient scrutiny. In automated test engineering, such deference can have concrete consequences, allowing flawed specifications or inadequate tests to shape the development of critical systems (Tiwari, 2025). In education, it can undermine academic integrity and the development of independent judgment, as students come to see AI as an unquestionable source of knowledge rather than a fallible and context bound assistant (Cotton et al., 2023; Chauncey and McKenna, 2023).

Ethical frameworks for AI in education and professional practice emphasize the need to preserve human agency, accountability, and fairness in the face of these challenges (Adams et al., 2023; Akgun and Greenhow, 2022). From this perspective, the goal is not to reject generative AI but to embed it within socio technical systems that promote critical engagement rather than passive consumption. Mixed initiative interaction, as articulated by Horvitz (1999), provides a valuable model here. In mixed initiative systems, control and decision making are dynamically shared between human and machine, with each party contributing according to its strengths. Applied to generative AI in BDD and education, this suggests designing workflows and interfaces that require users to review, modify, and justify AI outputs, thereby maintaining an active role in the reasoning process.

Practical examples of such scaffolding can be found in tools that highlight potential biases, uncertainties, or alternative interpretations in AI generated content (Harrison et al., 2024; Gordon et al., 2023). In writing support systems, for instance, prompts that encourage users to reflect on their goals, audience, and evidence can engage higher order cognitive processes, aligning with Bloom's taxonomy and its emphasis on analysis, evaluation, and creation (Forehand, 2010; Huitt, 2011). In automated test engineering, similar principles could be applied by requiring developers to articulate the rationale behind each AI generated scenario, to compare it with human authored requirements, and to discuss its implications with stakeholders.

The literature on digitally scaffolded debate and mobile blended interaction further supports this approach, showing that structured technological interventions can foster critical thinking and collaborative reasoning when they are designed to prompt reflection and dialogue rather than to provide definitive answers (Holzer et al., 2015; Holzer et al., 2018). These insights suggest that the risks of cognitive offloading and metacognitive erosion associated with generative AI are not inevitable but depend on how the technology is integrated into social and educational practices.

Nevertheless, significant challenges remain. One of the most difficult issues is the alignment of generative AI systems with human values and domain specific standards. Because these systems are trained on large and heterogeneous datasets, they may reproduce biases, errors, or inappropriate norms that are not easily detectable by users (Bender et al., 2021; Harrison et al., 2024). In behavior driven development, this could manifest as tests that reflect outdated practices, implicit assumptions about users, or incomplete coverage of critical scenarios, thereby embedding problematic perspectives into the software itself (Tiwari, 2025). In education, it could lead to the reinforcement of stereotypes or the marginalization of certain voices, undermining the ethical goals of equity and inclusion (Adams et al., 2023; Akgun and Greenhow, 2022).

Addressing these issues requires not only technical solutions but also institutional and pedagogical innovation. Teachers, developers, and organizations must cultivate a culture of critical AI literacy in which users understand the limitations, assumptions, and potential impacts of generative systems. This includes training in how to evaluate AI outputs, how to cross check information, and how to integrate human judgment with machine assistance in a responsible way (Chauncey and McKenna, 2023; Gordon et al., 2023). It also involves rethinking assessment and accountability structures so that they reward reflective engagement and ethical reasoning rather than mere efficiency or output quantity (Cotton et al., 2023; Facione et al., 1995).

Future research should build on this integrative framework by empirically examining how different designs of generative AI systems and workflows affect critical thinking, trust, and learning outcomes in both educational and software engineering contexts. Longitudinal studies could explore whether sustained use of AI assisted BDD leads to changes in developers' problem solving strategies, while classroom based experiments could test how different forms of AI scaffolding influence students' metacognitive awareness and transfer of knowledge. Such research would provide the evidence base needed to move from theoretical insight to practical guidance in the governance of generative artificial intelligence.

CONCLUSION

Generative artificial intelligence stands at a pivotal intersection of technological innovation, educational transformation, and cognitive change. Through its capacity to generate language, code, and structured artifacts, it has become a powerful force in domains ranging from classroom instruction to automated

behavior driven development. This study has argued that the significance of this force lies not only in its efficiency gains but in its profound impact on how humans think, learn, and collaborate with machines. By synthesizing research on critical thinking, human AI interaction, and test automation, the article has shown that generative AI can both enhance and erode the epistemic foundations of education and professional practice, depending on how it is designed and used.

The automation of behavior driven development, as articulated in recent scholarship, exemplifies this duality. It offers unprecedented opportunities to improve software quality and alignment with stakeholder needs, yet it also risks distancing developers from the reasoning processes that give those artifacts meaning and ethical grounding (Tiwari, 2025). Similarly, in education, generative AI can democratize access to knowledge and support, while simultaneously challenging the cultivation of independent judgment and reflective understanding that define true learning (Facione et al., 2011; Ahmad et al., 2023).

The path forward, therefore, lies not in rejecting generative AI but in integrating it within mixed initiative, ethically informed, and pedagogically sound systems that preserve human agency and critical engagement. By aligning technological automation with the goals of critical inquiry and responsible innovation, educators and engineers can harness the power of generative artificial intelligence while safeguarding the cognitive and moral capacities that make that power meaningful.

REFERENCES

1. Abunaseer, H. The use of generative AI in education: applications, and impact. Pressbooks.
2. Glikson, E., and Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals.
3. Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction.
4. Tiwari, S. K. (2025). Automating behavior driven development with generative AI: Enhancing efficiency in test automation. Frontiers in Emerging Computer Science and Information Technology.
5. Cotton, D. R. E., Cotton, P. A., and Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International.
6. Bender, J., Gebru, T., McMillan Major, A., and Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big.
7. Celik, I., Dindar, M., Muukkonen, H., and Jarvela, S. (2022). The promises and challenges of artificial intelligence for teachers.
8. Holzer, A., Govaerts, S., Bendahan, S., and Gillet, D. (2015). Towards mobile blended interaction fostering critical thinking.
9. Fernandes, D., Villa, S., Nicholls, S., Haavisto, O., Buschek, D., Schmidt, A., Kosch, T., Shen, C., and Welsch, R. (2024). AI makes you smarter, but none the wiser: The disconnect between performance and metacognition.
10. Adams, C., Pente, P., Lemermeyer, G., and Rockwell, G. (2023). Ethical principles for artificial intelligence in K 12 education.
11. Horvitz, E. (1999). Principles of mixed initiative user interfaces.
12. Facione, P. A., Sanchez, C. A., Facione, N. C., and Gainen, J. (1995). The disposition toward critical thinking.
13. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., and Amodei, D. (2020). Language models are few shot learners.
14. Gordon, A. D., Negreanu, C., Cambronero, J., Chakravarthy, R., Drosos, I., Fang, H., Mitra, B., Richardson, H., Sarkar, A., Simmons, S., Williams, J., and Zorn, B. (2023). Co audit: tools to help humans double check AI generated content.
15. Akgun, S., and Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K 12 settings.
16. Harrison, G., Bryson, K., Bamba, A. E., Dovichi, L., Binion, A. H., Borem, A., and Ur, B. (2024). JupyterLab in retrograde: Contextual notifications that highlight fairness and bias issues for data scientists.
17. Facione, P. A., Facione, N. C., and Sanchez, C. A. (1994). Critical thinking disposition as a measure of competent clinical judgment.
18. Ahmad, S. F., et al. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education.
19. Forehand, M. (2010). Blooms taxonomy.
20. Holzer, A., Tintarev, N., Bendahan, S., Kocher, B., Greenup, S., and Gillet, D. (2018). Digitally scaffolding debate in the classroom.
21. Ghosh, B., Narain, K., Guan, L., and Wilson, J. (2023). AI for everyone.
22. Chauncey, S. A., and McKenna, H. P. (2023). A framework and exemplars for ethical and responsible use of AI chatbot technology.

23. Greenwood, C., and Quinn, M. (2017). Digital amnesia and the future tourist.

24. Clark, J. H., Koenecke, A., Naecker, J., and Thorne, J. (2021). The role of artificial intelligence in enhancing education.

25. Huitt, W. (2011). Blooms taxonomy of the cognitive domain.

26. Hauser, D. J., and Schwarz, N. (2015). Its a trap: Instructional manipulation checks prompt systematic thinking on tricky tasks.