Report 1a

Manuel Oliveira

Human Technology Interaction Group

Department of Industrial Engineering & Innovation Sciences, TU/e

Project: Scaffolding writing skills using automated essay generation systems

Team: Rianne Conijn, Carlos Zednik, Bert Sadowski, Gunter Bombaerts

Funding: 4TU.CEE & BOOST!

Eindhoven, 23 December, 2024

Contents

Management summary 3
Context 3
Key challenges identified 3
Strategic responses and recommendations 3
Value for teachers and students 4
Conclusion and next steps 4
Introduction 5
1 - What future-oriented learning objectives are being identified as crucial in an AI-integrated educational landscape? 6
Brief commentary on the pace of GenAI development 7
GenAI versus Human abilities 8
On Big Tech, Power Dynamics and Structural Considerations 11
2 - How are educators transforming their assessment practices and developing new evaluation frameworks in response to GenAI technologies? 11
Practical strategies for redesigning assessments 14
Some recommendations 16
Learning Activities with AI 16
Discussion 17
Theoretical Implications and Future Directions 18
Next steps 19
References 20

Management summary

Context

Generative AI (GenAI) technologies like ChatGPT and similar chatbots are transforming higher education. As these tools become more sophisticated, they pose significant challenges and opportunities for teaching, learning, and assessment. In this summary, we provide an overview of the key insights from our more extensive report (available below), highlighting how our ongoing review of the state-of-the-art knowledge on the challenges and opportunities of GenAI in Higher Education can be translated into actionable information relevant to teachers and students, highlighting strategic directions, future research, and practical recommendations. The emphasis of this summary is primarily on practical recommendations, given the urgency to respond to the ongoing transformative impact of GenAI.

Key challenges identified

- **Disruption of traditional learning and assessment:** GenAI's humanlike content generation challenges conventional teaching methods, assessment integrity, and the unique value of human instruction (Kolade et al., 2024; Rathi et al., 2024).
- **Need for skill reorientation:** Both educators and students must shift focus toward skills that AI cannot easily replicate—critical thinking, creativity, ethical reasoning, adaptability, and AI literacy (Bower et al., 2024; Chauncey & McKenna, 2024; Kolade et al., 2024).
- **Assessment uncertainty:** With AI's capacity to generate content, verifying authorship and evaluating genuine understanding become more complex, necessitating a redesign of assessment methodologies (Fleckenstein et al., 2024; Jakesch et al., 2019).

Strategic responses and recommendations

Redefining learning objectives

- Emphasize higher-order skills: critical thinking, problem-solving, creative ideation, ethical reasoning, and adaptability (Kasneci et al., 2023; Zhai, 2022).
- Integrate AI literacy: Ensure students and teachers understand AI's capabilities, limitations, biases, and ethical considerations (Bower et al., 2024; Chiu, 2024; Kolade et al., 2024).

Transforming assessment practices

- Move from product-focused to process-oriented assessment: Evaluate reasoning processes, metacognitive skills, and real-world application (Cheng et al., 2024; Kolade et al., 2024).

- Adopt diverse, authentic evaluation methods: Use live presentations, peer assessments, project-based assignments, and frequent low-stakes assessments to mitigate AI's advantages in generating generic responses (Kolade et al., 2024; Xia et al., 2024).
- Leverage prompt analytics: Analyze student interactions with AI to gain insights into learning processes and provide personalized feedback (Cheng et al., 2024; M. Kim et al., 2024).

Faculty development and policy updates

- Invest in professional development to equip educators with skills for integrating AI responsibly into pedagogy and assessment (Chan & Tsi, 2024; Lim et al., 2023).
- Update assessment policies to establish clear guidelines on AI use, maintaining academic integrity while embracing AI-enabled learning enhancements (Mollick & Mollick, 2022; Xia et al., 2024).

Ongoing research and pilot studies

- Supporting and monitoring pilot projects (e.g., at TU/e) that assess the impact of GenAI on learning outcomes, teacher effectiveness, and student motivation.
- Continuous multidisciplinary research to refine teaching and assessment strategies, striving for an alignment with evolving (Gen)AI capabilities and future-oriented educational goals (X. Deng & Joshi, 2024; Mollick & Mollick, 2024; Rowland, 2023).

Value for teachers and students

For teachers:

- Empowerment through professional development and clear guidelines on AI integration, enabling them to design engaging, authentic assessments that emphasize unique human skills.
- Improved assessment tools and strategies that provide more accurate measures of student understanding and skill acquisition.

For students:

- Development of relevant, future-oriented skills that enhance employability and adaptability in an AI-driven landscape (Chiu, 2024; Zhai, 2022).
- Learning experiences that promote creativity, critical thinking, and ethical reasoning—areas where human judgment remains indispensable (Bower et al., 2024).

Conclusion and next steps

The integration of GenAI in higher education calls for a strategic, research-backed approach to curriculum design, assessment methods, and faculty development. By focusing on uniquely human skills and transforming assessment practices, institutions can harness AI's potential while preserving academic integrity and enhancing learning outcomes. Stakeholders are encouraged to support ongoing research, pilot projects, and policy updates that inform best practices. This proactive approach ensures that both teachers and students are prepared for and can thrive in an AI-integrated educational environment.

Introduction

The rapid advancement and increasing accessibility of generative AI (GenAI) technologies, and in particular ChatGPT (OpenAI, 2022) and its competitors (e.g., Anthropic's Claude, Google's Gemini, Meta's Llama; First Page Sage, 2024), is transforming the landscape of higher education. As these tools become more sophisticated and widely adopted (Hu, 2023), they present both opportunities and challenges for educators and students alike (Cong-Lem et al., 2024; Memarian & Doleck, 2023; Michel-Villarreal et al., 2023; Rahman & Watanobe, 2023; Ray, 2023; Stokel-Walker, 2022) . The ability of GenAI to produce humanlike content, solve complex problems, and engage in natural language interactions is disrupting traditional models of teaching, learning, and assessment (e.g., Fleckenstein et al., 2024; Kelly, 2023; Rashidi et al., 2023). This reality calls for a critical re-examination and redefinition of educational objectives, focusing on the skills and competencies students will need in an AI-driven world. The disruptive impact that GenAI tools is exerting on Education calls for urgent transformations in educational practices and learning objectives. Higher education institutions are consequently pressured to innovate and adapt their curricula, pedagogical approaches, and assessment strategies to align with this new reality while maintaining the integrity and value of the educational experience.

In a world where technology is increasingly capable of mimicking human ability by producing output that is undistinguishable from human-generated output (Fleckenstein et al., 2024; Gao et al., 2023; Kumar & Mindzak, 2024; J. Y. Lee, 2023; V. R. Lee et al., 2024) there is an increasing pressure for educators (and education institutions) to explain to their prospective students what are the benefits of learning from a human teacher, ideally supported by evidence-based arguments. At the same time, students are increasingly pressured to convince their future employers that their skill set remains valuable against agentic technology like GenAI or any of its successors. To answer such questions, we need to take steps to identify which (human) skills GenAI technology is less capable of replacing, and which new skills should both educators and students need to start developing to adapt to the age of GenAI.

Another important challenge that educators are facing in the age of GenAI relates to the assessment of learning and skill acquisition. One of the main features of content generation technologies is their ability to generate content that is not only humanlike (Fleckenstein et al., 2024; Gao et al., 2023; Kumar & Mindzak, 2024), but sometimes judged as more real than real content (Rathi et al., 2024; Tucciarelli et al., 2022). This type of technology directly impacts the ability of teachers to assess student learning in cases where the assessed content can be easily generated by GenAI tools, as the mere technological possibility creates a constant uncertainty about content authorship. Although GenAI tools threaten the effectiveness of many teaching activities like multiple-choice quizzes, the uncertainty of authorship is especially evident when student output is in the form of written content like reports or computer code (Banić et al., 2023; Gao et al., 2023; Groothuijsen et al., 2024; Kumar & Mindzak, 2024). Given the evidence demonstrating the low ability of humans to reliably detect cues of AI origin in written content (Fleckenstein et al., 2024; Jakesch et al., 2023), the stronger the need becomes for rethinking how student learning can be assessed in a manner that is compatible with the possibility (or perhaps requirement) of content created collaboratively with GenAI tools. It is therefore also crucial to understand what assessment methods will remain capable of supporting and ensuring the integrity and value

of learning, in an environment where the authorship of content is increasingly a mixture of machine and human inputs.

In this report we compile initial insights from an ongoing review of the literature aiming to address the following two questions branching from the challenges discussed above:

1) What future-oriented learning objectives are being identified as crucial in an AI-integrated educational landscape?

2) How are educators transforming their assessment practices and developing new evaluation frameworks in response to GenAI technologies?

Below, we provide a summary and analysis of the review results, and also address the closely related topic of learning activities, which is intricately connected with the other questions.

1 - What future-oriented learning objectives are being identified as crucial in an AI-integrated educational landscape?

In light of the pervasive influence of GenAI, higher education must prioritize the development of a range of cognitive, social, emotional, and technical competencies that will enable students to navigate and succeed in an AI-infused society. The recent literature emphasizes the importance of fostering skills such as critical thinking, problem-solving, creativity, and collaboration (Kasneci et al., 2023; Zhai, 2022). These higher-order thinking skills are essential for students to effectively leverage AI tools while maintaining the ability to think independently and generate original ideas. Additionally, digital literacy and AI literacy have emerged as crucial competencies, as students must learn to use AI tools ethically and responsibly, understanding their capabilities, limitations, and potential biases (Chiu, 2024; Zhai, 2022). Adaptability skills are also crucial in an environment where students are required to cope with and adapt to the continuously evolving landscape of AI-driven content generation technology (e.g., Bonfield et al., 2020; Chiu, 2024). By focusing on these future-oriented learning objectives and skills, higher education can empower students to thrive in a rapidly evolving technological landscape.

Below, we list the skills that were most frequently identified as relevant for the age of GenAI, followed by the skills that lost relevance as a result of the rise of GenAI.

The essential skills for the age of GenAI include:

- **AI Literacy:** Understanding how GenAI and AI in general work, along with its capabilities, limitations, and ethical issues (Bower et al., 2024; Chiu, 2024; Kolade et al., 2024; Ng et al., 2023).

- **Higher-order thinking:** Critical thinking, problem-solving, creativity, and analytical skills are crucial in an AI-driven world (Bower et al., 2024; Chauncey & McKenna, 2024; Kolade et al., 2024).
- **Learning with AI:** Using AI tools for learning, research, and problem-solving is vital (Bower et al., 2024; Chiu, 2024; Mollick & Mollick, 2022).

- **Adaptability:** The ability to embrace new technologies and changing work environments (Chauncey & McKenna, 2024; Chiu, 2024).

Skills becoming less relevant after the rise of GenAI include:

- **Knowledge Recall:** Simple information retrieval and memorization become less critical as AI handles these tasks (Bower et al., 2024; Kolade et al., 2024).

- **Routine Tasks:** Repetitive tasks like basic content generation or data entry can be automated (Kolade et al., 2024).

(Note: Please note this work is ongoing and this list might suffer alterations as the review progresses over time)

The literature review regarding the discussion of which skills will remain relevant in the age of GenAI proposes a shift from knowledge-based skills to those requiring human judgment, creativity, and critical thinking when working with AI tolls. Below we will also learn how the literature focused on the transformation of assessment practices to accommodate the new reality of GenAI technology is suggesting to focus on the evaluation of the processes involved in the acquisition and application of knowledge, as opposed to the evaluation of the output of knowledge application (e.g., Kolade et al., 2024).

Brief commentary on the pace of GenAI development

As of the submission date of this report, developers continue to release AI models claiming record achievements in large language model (LLM) based "reasoning," which may be more appropriately described as a form of *ersatz* or surrogate reasoning. An example is the OpenAI o3 family of models, announced on December 20th, 2024 (Zeff & Wiggers, 2024). Whether or not that is the case, the capacity of this new model to produce output that can only be verified by a minority of experts is likely to amplify the impact of this technology on education. Some potential effects of the continuous increase in the capabilities of these models to mimic reasoning (whether or not they indeed reason, e.g., Amirizaniani et al., 2024; Kambhampati, 2024) might manifest as an increased possibility of students outsourcing of tasks requiring complex reasoning to GenAI, posing a constant threat to assessment practices, or at least those which are incompatible with GenAI (e.g., blended

learning related assignments). Thus, one important question emerging in this discussion is also:

What are the human skills that GenAI is less capable or incapable of mimicking?

GenAI versus Human abilities

A quick search in the domain of grey literature provides us with some initial insights. In his article "Deep Learning Is Hitting a Wall," Gary Marcus (2022) argues that while LLMs are impressive in their ability to mimic human language, they are also fundamentally limited in their capacity for genuine understanding, reasoning, and truthfulness. Marcus (2022) emphasizes how LLMs struggle with common sense, logical inference, or handling novel situations, often producing nonsensical or even harmful outputs. These limitations result from their reliance on statistical patterns in immensely large datasets instead of on a true comprehension of the world, thereby rendering them unreliable in high-stakes scenarios (e.g. flying a plane, hiring decision, project planning This critique aligns with the concerns raised by Bender et al. (2021) in their seminal paper, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?", highlighting that these models, much like "stochastic parrots," can expertly mimic language without grasping its meaning. Bender and colleagues highlight not only the technical limitations but also the significant ethical and societal risks associated with the development and deployment of increasingly large and powerful LLMs.

In particular, Bender et al. (2021) point out that the sheer size and complexity of these models come with substantial environmental and financial costs, often concentrating power in the hands of a few resource-rich entities and potentially exacerbating existing inequalities. Moreover, they caution that the ability of LLMs to generate fluent and seemingly coherent text, while impressive, masks a profound lack of understanding. This can lead to the propagation of misinformation, the amplification of harmful biases, and the potential for malicious use, such as impersonation or the generation of deceptive content. Their work underscores the critical need for a more cautious approach that prioritizes ethical considerations, including the potential for social harm, alongside technical advancements. In essence, while Marcus (2022) focuses on the cognitive limitations of LLMs in terms of reasoning and understanding, Bender et al. (2021) broaden their critique to encompass the broader ethical and societal dimensions of deploying these technologies.

Together, these insights suggest that the path forward for (Gen)AI is not simply about building bigger and more powerful models but about developing systems that are more aligned with human values, capable of genuine understanding, and deployed responsibly within society (e.g. we can see glimpses of this latter behavior in the gradual release that most of these models undergo at launch in the past year). This perspective calls for a multidisciplinary approach that incorporates insights from computer science, linguistics, ethics, and the social sciences to minimize the disruption that AI development might cause otherwise, not only in education but to society as a whole.

The fundamental limitations of LLMs highlighted by Marcus (2022) and Bender et al. (2021) are further substantiated by recent empirical research. Amirizaniani et al. (2024) conducted a systematic investigation of LLMs' Theory of Mind (ToM) reasoning capabilities in open-ended scenarios, finding significant disparities between human and LLM reasoning processes. Their research demonstrates that even advanced models like GPT-4 struggle with nuanced social reasoning and complex open-ended questions, despite their impressive

performance on structured tasks. This aligns with Kambhampati's (2024) view that LLMs essentially perform "approximate retrieval" rather than genuine reasoning, while emphasizing that their ability to generate coherent text should not be mistaken for true understanding or principled reasoning.

The results of this review, supported by empirical research conducted by Amirizaniani et al., (2024) and theoretical insights by Kambhampati (2024), offer compelling evidence for the transformative impact of GenAI on assessment practices in higher education. While large language models (LLMs) demonstrate impressive capabilities in generating human-like text, their fundamental limitations in reasoning, understanding, and alignment with human values necessitate a critical re-evaluation of learning objectives and assessment methodologies. The research of Amirizaniani et al. (2024) highlights the significant disparities between human and LLM reasoning processes, particularly in open-ended scenarios that require nuanced social cognition and complex reasoning. This aligns with Kambhampati's (2024) characterization of LLMs as performing "approximate retrieval" rather than genuine reasoning. These findings underscore the need for a multifaceted approach to assessment that recognizes both the potential and the limitations of generative AI technologies.

Kambhampati's (2024) analysis further suggests that the distinction between human and AI capabilities lies not just in *what* tasks can be performed, but in *how* they are performed. This has important implications for assessment design:

- **Process-oriented assessment*:*** rather than focusing solely on final outputs, assessments should evaluate the reasoning process and methodology students employ.
- **Metacognitive skills**: assessment should target students' ability to reflect on and explain their thinking processes, as Amirizaniani et al.'s (2024) research shows how LLMs struggle to fully incorporate human emotions and intentions into their reasoning processes.
- **Contextual understanding**: Tasks should require students to demonstrate understanding across different contexts, as LLMs struggle with true transfer of knowledge despite their surface-level adaptability.

While LLMs excel at pattern matching and text generation based on statistical regularities in their training data (Bender et al., 2021), humans demonstrate empirically documented unique capabilities in complex social cognition, cultural learning, and theory of mind development (Laland & Seed, 2021) LLMs notably lack capabilities in establishing the correctness of their outputs from first principles (Kambhampati, 2024) and true understanding of communicative intent (Bender et al., 2021). Educational design should focus on developing these distinctly human capabilities, particularly in areas of collaborative learning, cultural knowledge transmission, and social understanding.

Table 1. Comparison of LLM Capabilities and Uniquely Human Skills

Skills/Competencies	Current LLM Capabilities	Uniquely Human Skills	Educational Implications	References
Information Processing	Pattern matching and text generation based on training data	Critical evaluation of information accuracy and relevance	Emphasis on information literacy and critical thinking skills	Bender et al. (2021) [Peer-reviewed article] Kambhampati (2024) [Peer-reviewed article]
Reasoning	Surface-level logical operations; struggles with novel situations	Deep causal reasoning, theory-building, and principled correctness assessment	Focus on complex problem-solving and knowledge integration across domains	Amirizaniani et al. (2024) [Peer-reviewed article] Kambhampati (2024) [Peer-reviewed article] Marcus (2022) [Opinion piece]
Social Understanding	Mimicry of social interactions; lacks true Theory of Mind	Genuine understanding of mental states, intentions, and communicative intent	Development of authentic social interaction and emotional intelligence skills	Amirizaniani et al. (2024) [Peer-reviewed article] Laland & Seed (2021) [Peer-reviewed article]
Creative Expression	Generation of variations based on training data	Original ideation, novel concept synthesis, and contextual adaptability	Assessment of unique perspectives, innovative thinking, and knowledge transfer	Bender et al. (2021) [Peer-reviewed article] Marcus (2022) [Opinion piece]
Ethical Judgment	Reproduction of ethical frameworks; lacks moral understanding	Nuanced moral reasoning, value-based decision making, and accountability	Integration of ethical reasoning components and responsible AI use in assessment tasks	Bender et al. (2021) [Peer-reviewed article] Kambhampati (2024) [Peer-reviewed article]

The comparative analysis of LLM capabilities and human skills presented in Table 1 provides a framework for educators (and institutions) to guide their assessment practices and curriculum design in the age of GenAI. The educational implications emphasize the development of critical thinking, information literacy, authentic social interaction, innovative thinking, and responsible AI use, all of which are essential for students to navigate the challenges and opportunities posed by these emerging technologies. Furthermore, the current review highlights the importance of a multidisciplinary approach that integrates insights from computer science, linguistics, ethics, and the social sciences to ensure the responsible development and deployment of GenAI in education. As Bender et al. (2021) argue, the ethical and societal dimensions of these technologies must be prioritized alongside technical advancements to mitigate potential risks and promote equitable outcomes.

Future research should focus on developing innovative assessment strategies that align with the unique capabilities of human cognition, such as process-oriented evaluation (e.g. analysis of student-GenAI interactions through prompt analytics), metacognitive skill

assessment, and contextual understanding (e.g., through analysis of how students infuse their knowledge in interactions with technology). Additionally, investigating effective methods for cultivating AI literacy among students and educators, examining the evolving landscape of teacher assessment literacy, and incorporating diverse stakeholder perspectives into AI-aware educational policies are crucial areas for further exploration.

On Big Tech, Power Dynamics and Structural Considerations

The empirical findings regarding LLMs' cognitive limitations must be contextualized within the broader structural dynamics of AI development and deployment. Research conducted by Widder and colleagues (2024) reveals that the development of large language models requires massive computational resources, specialized hardware, and extensive datasets predominantly controlled by a small number of corporate entities. This creates structural *choke points* in AI development, where even well-resourced academic and research institutions struggle to operate independently of market channels controlled by major technology companies. The infrastructure requirements - from specialized AI chips (e.g., nVidia) to cloud computing platforms - further aggravate these dependencies, creating a systemic bias in how AI capabilities are developed, studied, and deployed.

These structural conditions have profound methodological implications for educational research and practice. Rather than democratizing access to advanced computational capabilities, the current architecture may exacerbate existing inequalities as educational institutions become increasingly dependent on corporate infrastructure to deploy AI tools. As Khanal et al. (2024) demonstrate through policy analysis, major technology companies have positioned themselves as "actors of immense strategic interest" (Khanal et al., 2024, p.13) wielding substantial influence over both technological development trajectories and institutional practices.

Together, the ongoing discussion suggests that research into educational AI applications must carefully consider not only the cognitive limitations of current systems, but also the *power dynamics and dependencies* that shape how these technologies can be studied, developed, and integrated into educational contexts.

2 - How are educators transforming their assessment practices and developing new evaluation frameworks in response to GenAI technologies?

Traditional assessment methods, such as essays and exams, are increasingly vulnerable to the disruptive effects of GenAI, which can easily produce content that mimics human-generated responses (Kolade et al., 2024; Rathi et al., 2024; Xia et al., 2024). For instance, Xia and colleagues (2024) emphasize that GenAI challenges the validity of traditional assessments such as essays by enabling students to generate high-quality content effortlessly. The ability to outsource cognitive effort to GenAI tools has been shown to disrupt

the learning experience (Bastani et al., 2024) and content co-created with AI is perceived by the authors as less meaningful to them and to others (Campbell et al., 2024). On a similar vein, Kolade and collaborators (2024) highlight the limitations of existing summative assessments in detecting AI-generated content and propose shifting the focus from assessing knowledge reproduction to assessments of how students apply their knowledge and competencies. To maintain the integrity and effectiveness of assessments in an AI-infused educational environment, educators are therefore urged to adopt innovative strategies that minimize the impact of (Gen)AI-generated content and that promote a more authentic learning (Cheng et al., 2024; J. Kim et al., 2022; Xia et al., 2024). This technological shift demands a fundamental rethinking of assessment strategies to both maintain academic integrity and better evaluate students' actual learning and capabilities (Kolade et al., 2024; Xia et al., 2024). Rather than viewing (Gen)AI as a purely disruptive technology, educators can alternatively perceive it as an opportunity to transform their educational philosophies and practices and take steps to implement more authentic, process-focused assessment methods that can better prepare students for the age of AI (Cheng et al., 2024; Mollick & Mollick, 2022; Xia et al., 2024).

Our current review of the literature highlights some themes and recommendations regarding how educators can adapt their assessment methods to deal with the reality of GenAI in the classroom. We list them below in no particular order of importance.

Impact of GenAI on assessment

GenAI is fundamentally challenging traditional assessment methods. The ease with which AI can generate human-quality text necessitates a rethinking of assessment design to ensure academic integrity and accurately evaluate student learning (Kolade et al., 2024; Xia et al., 2024).

Detection of AI-generated content

While some researchers are exploring methods to detect AI-generated text (Guo et al., 2023; Kumar & Mindzak, 2024; Liang et al., 2023; Sardinha, 2024),the rapid evolution of AI models makes it increasingly challenging. A shift from focusing solely on detection to designing assessments that are less susceptible to AI assistance is crucial, especially in light of growing evidence showing how humans are unable to realiabley and flawlessly distinguish between content generated by humans or AI (Fleckenstein et al., 2024; Jakesch et al., 2023).

Collaboration, Not Replacement

(Gen)AI is not meant to replace educators but rather to augment their capabilities. The focus should be on how AI can assist teachers in providing more personalized support and fostering deeper student engagement (Chan & Tsi, 2024; Lim et al., 2023).

New possibilities for learning

GenAI tools carry the potential to augment student learning, particularly in areas like writing, planning and collaborative problem-solving. They can provide personalized feedback at any point in time, support self-assessment, and facilitate deeper understanding through interactive dialogue (e.g., Cheng et al., 2024; Kasneci et al., 2023; Memarian & Doleck, 2023; Mollick & Mollick, 2022).

Ethical considerations

The rise of AI raises ethical concerns, including academic integrity, potential biases in AI outputs, and the need for transparent guidelines for AI use in educational settings (X. Deng & Joshi, 2024; Rowland, 2023). The reviewed literature stresses the need to address these challenges for a responsible and fair AI integration in the context of education (Mollick & Mollick, 2022; Xia et al.).

Prompt engineering and chatbot interaction analysis

The way students interact with AI tools through prompts reveals valuable insights into their learning processes. Analyzing prompt data (what one may refer to as "prompt analytics") can provide educators with information to personalize feedback, identify areas of difficulty, and tailor instructional strategies (Cheng et al., 2024; J. Kim et al., 2022).

Re-designing assessment

The reviewed literature highlights the need to move beyond traditional essay-based assessments to start exploring alternative formats such as: oral exams, project-based assignments, or assessments that emphasize critical thinking and problem-solving skills, which are more difficult for AI to replicate (Kolade et al., 2024). Moreover, it recommends the investment on learning activities that promote GenAI literacy in both teachers and students, with the aim to find novel ways to assess knowledge through the analysis of interactions with GenAI tools (Mollick & Mollick, 2022, 2023, 2024).

Emphasis on process and metacognition

The integration of GenAI in educational activities allows for a greater emphasis on the learning process itself. Encouraging students to reflect on their interactions with GenAI, analyze its feedback, and articulate their thought processes can improve metacognitive skills and deeper learning (Cheng et al., 2024; Garrison et al., 1999)

The current summary of insights derived from the literature reveals several interconnected themes: the need to move beyond detection-focused approaches, the importance of viewing AI as an augmentative rather than replacement technology, and the emergence of novel opportunities for learning through AI-human interaction. The literature consistently emphasizes shifting focus from product to process, particularly through analysis of student-AI interactions and metacognitive development. These findings suggest that successful adaptation to an AI-integrated educational landscape requires a multi-faceted approach that combines redesigned assessment strategies, enhanced AI literacy, and careful

consideration of ethical implications. This transformation presents both significant challenges and opportunities for innovation in educational assessment, pointing towards a future where technology and pedagogy can be meaningfully integrated to enhance learning outcomes while maintaining academic integrity. The reviewed contents indicate that by embracing these changes thoughtfully and systematically, universities can leverage the potential of Genal while at the same time preserving he fundamental values of higher education.

Practical strategies for redesigning assessments

The reality of GenAI requires not only an integration of the collective experiences of educators worldwide, but also clear solutions that educators can rapidly adopt in their response to the rise of GenAI in the classroom. Below we provide a range of strategies that are currently being explored in the context of higher education (Charles Sturt University, 2024). These strategies can serve as a source of inspiration for educators interested in redesigning their assessment methods to ensure these are resilient to GenAI's capabilities. These include:

- **Incorporate authentic assessment**: Authentic assessments require students to apply their knowledge and skills to real-world scenarios. These tasks often involve complex problem-solving, critical analysis, and the ability to adapt knowledge to new situations. While GenAI can generate text based on existing information, it struggles with the unpredictable and nuanced nature of authentic tasks that require genuine understanding and creative application of knowledge. Examples of authentic assessment include case studies, simulations, and projects that address real-world issues.

- **Promote critical thinking**: Assessments designed to evaluate critical thinking skills require students to analyze information from multiple sources, identify biases, construct logical arguments, and draw well-supported conclusions. These are higher-order cognitive skills that current GenAI systems struggle to replicate convincingly. Focusing on critical thinking ensures students are not simply regurgitating information but engaging deeply with the subject matter and developing their analytical abilities.

- **Individualized or personalized assessments**: Tailoring assessments to individual student needs, learning styles, and interests creates a more personalized learning experience that is difficult for GenAI to mimic. This can involve offering students choices in assessment formats, adjusting the complexity of tasks, or incorporating personal interests into project topics. Personalized assessments move away from standardized, one-size-fits-all approaches, making it harder for GenAI to generate generic responses that fit every student.

- **Live or recorded interviews or presentations**: Live or recorded interviews and presentations allow educators to assess a student's understanding and communication skills in real-time. The spontaneous and interactive nature of these formats makes it challenging for GenAI to generate appropriate responses on the spot. This approach emphasizes the human element of communication, adaptability, and critical thinking in unscripted situations.

- **Find novel or unique ways of designing application-based questions**: By designing assessment questions that require creative and innovative application of knowledge, educators can challenge the limitations of current GenAI systems. GenAI excels at pattern recognition and reproducing existing information, but it struggles with novel tasks that require originality, flexible thinking, and the ability to synthesize information in new ways. This strategy encourages students to think beyond formulating answers and demonstrate a deeper level of understanding.

- **Peer assessments**: Engaging students in peer assessment activities can provide valuable feedback and promote collaborative learning. While GenAI could potentially be used to generate feedback, peer assessment emphasizes the human element of evaluating work, providing constructive criticism, and learning from different perspectives. This approach also encourages students to reflect critically on their own work and develop a deeper understanding of assessment criteria.

- **Frequent low-stakes assessments**: Frequent low-stakes assessments provide opportunities for regular feedback and allow students to demonstrate their understanding throughout the course. The cumulative nature of these assessments makes it more difficult for GenAI to generate consistent, high-quality responses across multiple instances. This approach focuses on ongoing learning and progress, rather than relying solely on high-stakes exams that are more susceptible to GenAI assistance.

- **Promote creativity and problem-solving**: Assessments that emphasize creativity and problem-solving require students to generate original ideas, develop innovative solutions, and think critically about complex challenges. These are skills that are currently beyond the capabilities of GenAI systems, which rely on existing data and patterns to generate responses. By focusing on these higher-order cognitive skills, educators can assess learning in ways that are less susceptible to GenAI manipulation.

- **Integrate real-life situations and practical experiences**: Assessments that integrate real-life situations and practical experiences require students to apply their knowledge in meaningful contexts. GenAI can provide information, but it lacks the lived experience and contextual understanding needed to navigate the complexities of real-world scenarios. This approach allows educators to assess students' ability to transfer knowledge, adapt to new situations, and demonstrate practical competence.

Educators can incorporate these strategies in their assessment practices to mitigate the risks posed by GenAI, while promoting the skills essential for academic and professional success (addressed in the first research question).

Some recommendations

Our brief review of the literature allows us to offer some general recommendations to educators in the age of GenAI:

- **Faculty development:** Universities must provide comprehensive professional development opportunities for faculty to understand the capabilities and limitations of AI, address ethical considerations, and effectively integrate AI tools into their pedagogy and assessment practices.
- **Update assessment policies:** Universities need to revisit and revise their assessment policies to account for the use of AI. Clear guidelines on appropriate AI use, academic integrity, and expectations for student work are essential.
- **Focus on higher-order thinking skills:** Curricula and assessments should prioritize higher-order thinking skills, such as critical analysis, problem-solving, creativity, and ethical reasoning, which are less easily replicated by AI.
- **Need for continuous research:** Continuous research on the impact of AI in education is critical to inform best practices, understand the evolving nature of AI, and ensure that its integration serves to enhance learning and promote equitable access to quality education.

Learning Activities with AI

While this section is not directly related to any of the research questions guiding our literature review above, it is strongly related with both of them in the sense that, in comnbination with the learning objectives and assessment elements that are at the core of pedagogy, learning activities represent the means by which the learning objectives are promoted in learning. Together, these three elements should be aligned in a constructive triangle, the well-known principles of constructive alignment in education proposed by Biggs and Tang (2011).

The effective preparation of students for an AI-augmented professional landscape necessitates a methodological reconceptualization of learning activities in higher education (Mollick & Mollick, 2023, 2024; Xia et al., 2024). Our analysis of the literature indicates that

project-based and problem-based pedagogical frameworks offer particularly promising approaches for developing critical competencies through their emphasis on authentic task engagement and interdisciplinary collaboration (Essel et al., 2024). These methodologies facilitate the systematic development of higher-order cognitive skills while simultaneously providing structured opportunities for GenAI tool integration.

A strategic incorporation of GenAI technologies into educational activities can enhance learning outcomes through multiple mechanisms. AI-enabled adaptive learning systems show significant potential for optimizing instructional personalization through continuous calibration to individual learning trajectories (Kasneci et al., 2023). GenAI tools like ChatGPT and similar chatbots can be thoughtfully integrated into learning activities to increase student engagement and skill development. For instance, in knowledge-building activities, GenAI can serve as a springboard for idea generation and critical discussion (Chen et al., 2023). Students can use GenAI to generate initial drafts, explanations, or inquiry questions, and then collaboratively evaluate, refine, and build upon these AI-generated contributions. In programming courses, GenAI can be used for tasks like error checking, debugging, and code explanation, allowing students to focus on higher-order problem-solving skills (Groothuijsen et al., 2024). However, it is crucial to design learning activities that require students to go beyond simply using GenAI outputs (R. Deng et al., 2024). Project-based assessments and tasks that demand originality and the application of knowledge in unique contexts can help distinguish between AI assistance and genuine learning gains.

The current review leads us to the view that the strategic integration of AI technologies in education, when aligned with established pedagogical frameworks, will create learning environments that effectively prepare students for algorithmic-human professional contexts. This carries significant implications for how institutions should approach technology-enhanced learning while preserving core academic objectives.

Discussion

The emergence of GenAI technologies like ChatGPT or its contemporary competitors represents a transformative moment in higher education, challenging traditional paradigms of learning objectives and assessment practices across multiple disciplines (Chiu, 2024; Kasneci et al., 2023; Xia et al., 2024). In our review of the literature we examine the transformative impact of GenAI technologies on higher education, with particular focus on pedagogical adaptation and assessment methodology. Two primary research questions guided this investigation: (1) the identification of critical future-oriented learning objectives in GenAI-integrated educational environments, and (2) the development of effective assessment frameworks that accurately evaluate student competencies in the context of GenAI availability.

Regarding the identification of the most relevant future-oriented skills, our analysis of the literature reveals a systematic shift toward competencies that demonstrate distinctive human cognitive capabilities. Critical findings indicate the emergence of three primary domains:

- **Cognitive Processing Skills**: Research demonstrates the paramount importance of developing advanced cognitive capabilities, particularly in areas where AI systems show limitations. These include complex problem-solving frameworks, critical

analytical thinking, and creative synthesis of information (Bower et al., 2024; Chauncey & McKenna, 2024).

- **AI Literacy Competencies**: Empirical investigations highlight the necessity of developing sophisticated understanding of AI systems, encompassing both technical comprehension and ethical implications. This literacy framework enables students to engage critically with AI tools while maintaining awareness of systemic limitations.
- **Adaptive Learning Capabilities**: Studies indicate the critical nature of developing adaptive competencies, preparing students to navigate evolving technological landscapes while maintaining academic rigor.

Regarding the question of how assessment can be redesigned to accommodate GenAI technology in the context of higher education, our review suggests a significant methodological shift in assessment practices, characterized by:

- Integration of process-oriented evaluation methodologies
- Development of metacognitive assessment protocols
- Implementation of collaborative evaluation frameworks

Empirical evidence further suggests that current GenAI systems exhibit limitations in complex reasoning and contextual interpretation (Amirizaniani et al., 2024; Kambhampati, 2024), requiring assessment frameworks that emphasize: cognitive strategy evaluation, ethical reasoning assessment, and metacognitive process analysis,

Theoretical Implications and Future Directions

The synthesis of research findings indicates a fundamental epistemological transformation in educational theory. Studies consistently demonstrate the emergence of a collaborative model where GenAI functions as an augmentative tool rather than a replacement mechanism (Chan & Tsi, 2024; Mollick & Mollick, 2022). This theoretical framework emphasizes: the integration of authentic assessment methodologies, the development of personalized learning protocols, and the implementation of project-based pedagogical frameworks. The research strongly indicates the need for a multidisciplinary theoretical approach, incorporating perspectives from computer science (i.e., knowledge about the inner workings of AI and its technical aspects), linguistics, ethics, and social sciences (X. Deng & Joshi, 2024; Rowland, 2023). This integrated framework facilitates the responsible integration of (Gen)AI technologies while also maintaining focus on fundamental educational objectives and societal values.

The current (and ongoing) review contributes to the theoretical understanding of (Gen)AI integration in higher education while additionally highlighting critical areas for future (ideally more empirically-oriented) investigation. The findings suggest the need for continued research into assessment methodologies that effectively evaluate human cognitive capabilities while acknowledging the evolving role of GenAI technologies in educational contexts.

Next steps

As the landscape of GenAI continues to evolve, higher education institutions are called upon to engage in continuous research, develop adaptive pedagogical strategies, and commit to educational practices that augment rather than replace human cognitive potential.

At TU/e, our project will continue on parallel threads: On the first thread, we will continue reviewing the literature to track the fast-paced developments in the field of generative AI and its impact on higher education. On a second thread, the team will be involved in designing, implementing and evaluating pilot studies assessing the impact of GenAI on various domains ranging from student learning to teaching activities, or assessing the value of innovative assessment methods such as the assessment of learning through prompt analytics. On a third thread, the project is currently compiling and analyzing data from several pilot studies in TU/e courses spanning academic years from 2022 until the present (Data Science Ethics, Philosophy and Ethics of AI, Rational Agents), where teachers have been developing and implementing an assessment rubric to evaluate learning through students' interaction with GenAI tools. The current data will be augmented by the ongoing pilot studies in AY 2024-2025, and is expected to lead to a scientific publication in 2025 focused on providing empirical insights into the ability of our GenAI-compatible assessment approach to capture learning, clear recommendations for educators, and integrated insights derived from gathering students' perspectives on the evaluation of interactions with GenAI in the classroom context.

References

Amirizaniani, M., Martin, E., Sivachenko, M., Mashhadi, A., & Shah, C. (2024). Can LLMs Reason Like Humans? Assessing Theory of Mind Reasoning in LLMs for Open-Ended Questions. *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management*, 34–44. https://doi.org/10.1145/3627673.3679832

Banić, B., Konecki, M., & Konecki, M. (2023). Pair Programming Education Aided by ChatGPT. *2023 46th MIPRO ICT and Electronics Convention (MIPRO)*, 911–915. https://doi.org/10.23919/MIPRO57284.2023.10159727

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2024). *Generative AI can harm learning* (SSRN Scholarly Paper 4895486). https://doi.org/10.2139/ssrn.4895486

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? 🦜. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

Biggs, J., & Tang, C. S. (with Society for Research into Higher Education). (2011). *Teaching for quality learning at university: What the student does* (4th edition). McGraw-Hill, Society for Research into Higher Education & Open University Press.

Bonfield, C. A., Salter, M., Longmuir, A., Benson, M., & Adachi, C. (2020). Transformation or evolution?: Education 4.0, teaching and learning in the digital age. *Higher Education Pedagogies*, *5*(1), 223–246. https://doi.org/10.1080/23752696.2020.1816847

Bower, M., Torrington, J., Lai, J. W. M., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies*, *29*(12), 15403–15439. https://doi.org/10.1007/s10639-023-12405-0

Campbell, A. V., Wang, Y., & Inzlicht, M. (2024). *Experimental evidence that exerting effort increases meaning*. OSF. https://doi.org/10.31234/osf.io/z3g9c

Chan, C. K. Y., & Tsi, L. H. Y. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, *83*, 101395. https://doi.org/10.1016/j.stueduc.2024.101395

Charles Sturt University. (2024). *Rethinking assessments in the age of artificial intelligence*. Charles Sturt University. https://www.csu.edu.au/division/learning-teaching/assessments/assessment-and-artificial-intelligence/rethinking-assessments

Chauncey, S. A., & McKenna, H. P. (2024). Exploring the Potential of Cognitive Flexibility and Elaboration in Support of Curiosity, Interest, and Engagement in Designing AI-Rich Learning Spaces, Extensible to Urban Environments. In N. A. Streitz & S. Konomi (Eds.), *Distributed, Ambient and Pervasive Interactions* (pp. 209–230). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-60012-8_13

Chen, B., Zhu, X., & Díaz del Castillo H., F. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence*, *5*, 100184. https://doi.org/10.1016/j.caeai.2023.100184

Cheng, Y., Lyons, K., Chen, G., Gašević, D., & Swiecki, Z. (2024). Evidence-centered assessment for writing with generative AI. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 178–188. https://doi.org/10.1145/3636555.3636866

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, *6*, 100197. https://doi.org/10.1016/j.caeai.2023.100197

Cong-Lem, N., Soyoof, A., & Tsering, D. (2024). A systematic review of the limitations and associated opportunities of ChatGPT. *International Journal of Human–Computer Interaction*, 1–16. https://doi.org/10.1080/10447318.2024.2344142

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 105224. https://doi.org/10.1016/j.compedu.2024.105224

Deng, X., & Joshi, K. D. (2024). Promoting ethical use of generative ai in education. *SIGMIS Database*, *55*(3), 6–11. https://doi.org/10.1145/3685235.3685237

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, *6*, 100198. https://doi.org/10.1016/j.caeai.2023.100198

First Page Sage. (2024). *Top generative ai chatbots by market share—December 2024*. https://firstpagesage.com/reports/top-generative-ai-chatbots/

Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence*, *6*, 100209. https://doi.org/10.1016/j.caeai.2024.100209

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *Npj Digital Medicine*, *6*(1), 1–5. https://doi.org/10.1038/s41746-023-00819-6

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education*, *2*(2), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Groothuijsen, S., van den Beemt, A., Remmers, J. C., & van Meeuwen, L. W. (2024). AI chatbots in programming education: Students' use in a scientific computing course and consequences for learning. *Computers and Education: Artificial Intelligence*, *7*, 100290. https://doi.org/10.1016/j.caeai.2024.100290

Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J., & Wu, Y. (2023). *How close is ChatGPTto human experts? Comparison corpus, evaluation, and detection* (arXiv:2301.07597). arXiv. http://arxiv.org/abs/2301.07597

Hu, K. (2023, February 2). *ChatGPT sets record for fastest-growing user base—Analyst note*. Reuters. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Jakesch, M., French, M., Ma, X., Hancock, J. T., & Naaman, M. (2019). AI-Mediated Communication: How the Perception that Profile Text was Written by AI Affects Trustworthiness. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3290605.3300469

Jakesch, M., Hancock, J. T., & Naaman, M. (2023). Human heuristics for AI-generated language are flawed. *Proceedings of the National Academy of Sciences*, *120*(11), e2208839120. https://doi.org/10.1073/pnas.2208839120

Kambhampati, S. (2024). Can large language models reason and plan? *Annals of the New York Academy of Sciences*, *1534*(1), 15–18. https://doi.org/10.1111/nyas.15125

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (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

Kelly, M. L. (2023, January 26). “Everybody is cheating”: Why this teacher has adopted an open ChatGPT policy. *NPR*. https://www.npr.org/2023/01/26/1151499213/chatgpt-ai-education-cheating-classroom-wharton-school

Khanal, S., Zhang, H., & Taeihagh, A. (2024). Why and how is the power of Big Tech increasing in the policy process? The case of generative AI. *Policy and Society*, puae012. https://doi.org/10.1093/polsoc/puae012

Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, *27*(5), 6069–6104. https://doi.org/10.1007/s10639-021-10831-6

Kim, M., Kim, S., Lee, S., Yoon, Y., Myung, J., Yoo, H., Lim, H., Han, J., Kim, Y., Ahn, S.-Y., Kim, J., Oh, A., Hong, H., & Lee, T. Y. (2024). *Designing Prompt Analytics Dashboards to Analyze Student-ChatGPT Interactions in EFL Writing* (arXiv:2405.19691). arXiv. https://doi.org/10.48550/arXiv.2405.19691

Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon*, *10*(4), e25953. https://doi.org/10.1016/j.heliyon.2024.e25953

Kumar, R., & Mindzak, M. (2024). Who Wrote This? Detecting Artificial Intelligence–Generated Text from Human-Written Text. *Canadian Perspectives on Academic Integrity*, *7*(1), Article 1. https://doi.org/10.55016/ojs/cpai.v7i1.77675

Laland, K., & Seed, A. (2021). Understanding human cognitive uniqueness. *Annual Review of Psychology*, *72*(1), 689–716. https://doi.org/10.1146/annurev-psych-062220-051256

Lee, J. Y. (2023). Can an artificial intelligence chatbot be the author of a scholarly article? *Journal of Educational Evaluation for Health Professions*, *20*, 6. https://doi.org/10.3352/jeehp.2023.20.6

Lee, V. R., Pope, D., Miles, S., & Zárate, R. C. (2024). Cheating in the age of generative AI: A high school survey study of cheating behaviors before and after the release of ChatGPT. *Computers and Education: Artificial Intelligence*, *7*, 100253. https://doi.org/10.1016/j.caeai.2024.100253

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). *GPT detectors are biased against non-native English writers* (arXiv:2304.02819). arXiv. https://doi.org/10.48550/arXiv.2304.02819

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, *21*(2), 100790. https://doi.org/10.1016/j.ijme.2023.100790

Marcus, G. (2022). *Deep Learning Is Hitting a Wall*. https://nautil.us/deep-learning-is-hitting-a-wall-238440/

Memarian, B., & Doleck, T. (2023). ChatGPT in education: Methods, potentials, and limitations. *Computers in Human Behavior: Artificial Humans*, *1*(2), 100022. https://doi.org/10.1016/j.chbah.2023.100022

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *Education Sciences*, *13*(9), Article 9. https://doi.org/10.3390/educsci13090856

Mollick, E. R., & Mollick, L. (2022). *New Modes of Learning Enabled by AI Chatbots: Three Methods and Assignments* (SSRN Scholarly Paper 4300783). https://doi.org/10.2139/ssrn.4300783

Mollick, E. R., & Mollick, L. (2023). Using AI to Implement Effective Teaching Strategies in Classrooms: Five Strategies, Including Prompts. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4391243

Mollick, E. R., & Mollick, L. (2024). *Instructors as innovators: A future-focused approach to new AI learning opportunities, with prompts* (SSRN Scholarly Paper 4802463). https://doi.org/10.2139/ssrn.4802463

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, *71*(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

OpenAI. (2022). *ChatGPT (Nov 30 version) [Large language model]* [Computer software]. OpenAI. https://chat.openai.com/chat

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), Article 9. https://doi.org/10.3390/app13095783

Rashidi, H. H., Fennell, B. D., Albahra, S., Hu, B., & Gorbett, T. (2023). The ChatGPT conundrum: Human-generated scientific manuscripts misidentified as AI creations by AI text detection tool. *Journal of Pathology Informatics*, *14*, 100342. https://doi.org/10.1016/j.jpi.2023.100342

Rathi, I., Taylor, S., Bergen, B. K., & Jones, C. R. (2024). *GPT-4 is judged more human than humans in displaced and inverted Turing tests* (arXiv:2407.08853). arXiv. https://doi.org/10.48550/arXiv.2407.08853

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, *3*, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Rowland, D. R. (2023). Two frameworks to guide discussions around levels of acceptable use of generative AI in student academic research and writing. *Journal of Academic Language and Learning*, *17*(1), Article 1.

Sardinha, T. B. (2024). AI-generated vs human-authored texts: A multidimensional comparison. *Applied Corpus Linguistics*, *4*(1), 100083. https://doi.org/10.1016/j.acorp.2023.100083

Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—Should professors worry? *Nature*. https://doi.org/10.1038/d41586-022-04397-7

Tucciarelli, R., Vehar, N., Chandaria, S., & Tsakiris, M. (2022). On the realness of people who do not exist: The social processing of artificial faces. *iScience*, 105441. https://doi.org/10.1016/j.isci.2022.105441

Widder, D. G., Whittaker, M., & West, S. M. (2024). Why 'open' AI systems are actually closed, and why this matters. *Nature*, *635*(8040), 827–833. https://doi.org/10.1038/s41586-024-08141-1

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, *21*(1), 40. https://doi.org/10.1186/s41239-024-00468-z

Zeff, M., & Wiggers, K. (2024, December 20). OpenAI announces new o3 models. *Tech Crunch*. https://techcrunch.com/2024/12/20/openai-announces-new-o3-model/

Zhai, X. (2022). ChatGPT user experience: Implications for education. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4312418