International Journal of Education and Humanities
ISSN: 2770-6702 | Vol. 23, No. 2, 2026

Effects of Generative AI on Students' Cognitive Engagement from a Multimodal Learning Perspective

Zihan Yuan

Faculty of Education, The University of Melbourne, Parkville 3010, Melbourne, Australia

Abstract: The wide spread use of generative AIs in teaching may raise questions about whether, via technology itself, we might be able to foster even greater interest among students. Based on the above-mentioned theories including Mayer's theory of cognitively structured media in multimedia learning, Fredricks and his colleagues' three-component model of student engagement, and Sweller's cognitive load theory to explore how generative AI tools influence students' cognitive engagement from a perspective of multimodal learning theory; The main point is that Generative AI brings about a qualitative change in the environment for multimodal learning - Conditions such as dynamic personalised interactive content generation, including text-image-dialogue modes, which can enhance the quality of cognition while carrying potential dangers of shallow processing and learners' dependence on technology. The article first explores how generative artificial intelligence can trigger more profound cognitive processes; Then investigate the circumstances under which these mechanisms tend to enhance rather than replace true intellectual work, And Finally consider its consequences for pedagogy Construction design teachers Professional Development, The equitable distribution of use in diverse learning situations.

Keywords: Generative AI, Cognitive Engagement, Multimodal Learning, Cognitive Load, Large Language Models, Educational Technology, Self-Regulated Learning.

1. Introduction

Generative Artificial Intelligence is an umbrella term that includes all kinds of large-scale model systems capable of generating high-quality texts, images, codes, and other content when prompted by natural language; moreover, it can produce different types of output simultaneously in various modalities under such circumstances. GPT-4, which demonstrated its technical capability in documenting OpenAI in 2023, can provide explanation, answer questions, create working examples, simulate conversation partners, and adjust the output based on the learner's characteristics according to various aspects of this difference compared with previous versions of educational technology[4]. These abilities do not have a fixed level of positivity or negativity; rather, they change with the integration path of generative artificial intelligence tools in teaching activities and the requirements placed by such paths on learners' cognitive participation.

Cognitive engagement - the extent to which learners put in mental efforts when participating in learning activities; employs deeper forms of thinking than superficial ones; can regulate one's own learning independently autonomously - it has been confirmed by scholars that cognitive engagement stands apart from behavioural engagement or emotional engagement yet correlates positively with both [2]. Therefore, whether or not generative AI affects cognitive engagement has long been a problem in the research on educational-AI; If these tools increase users' actions without raising their level of cognitive struggle and thus lead to decreasing retention among students over time, then this type of tool might harm education. However, if AI tools are developed and implemented in a manner that limits deep cognitive processing, inhibits metacognitive awareness, or fails to provide productive struggles that lead to learning gains as per current education research findings, then their impact on improving teaching effectiveness will not be significant.

Based on a multimodal Learning perspective, This paper will provide an appropriate analysis Method To solve these Problems Detail And determine Which Combination Of Representations Is More beneficial In terms Of Perceptual Improvement. According to Mayer's cognitive theory, it is not necessary for the medium environment to be rich; what matters most is whether there is a match among different representations and people learn in their minds [1]. Generative AIs bring another layer of solution for alignment problems, that is, its outputs are generated dynamically as learners input data changes; In addition, generative AI can combine various media in an unexpected way not covered by all instructional materials before them. How the unique characteristics mentioned above interact with the cognitive mechanism identified by the literature on multimodal learning needs to be synthesised theoretically and verified empirically carefully.

2. Theoretical Foundations

Based on Mayer's cognitive theory of multimedia learning, this paper will explore the impact that generative AI-generated results have had on students'. Based on the following three fundamental laws of cognitive science: Firstly, people can only perceive verbi- graphical materials in two ways separately; Secondly, there is a limit to their processing capabilities for both types of information simultaneously; Thirdly, meaningful learning requires active integration among different pieces of information by learners. Based on this framework, Mayer obtained a series of Design rules - Coherence Principle , signalling Principle , Modality Principle to specify the Conditions Under Which Multimodal Instruction Supports Rather Than Overload Cognitive Processing[1]. The above-mentioned principles are based on the scenario of having a fixed teaching resource, however, they still have an essential relevance to the situation that generative AI provides dynamic content for learners; The risk of extra cognitive burden caused by unreasonably arranged or overly complicated multi-modal output environment exceeds

that under controlled environments.

Fredricks et al.'s three-part model breaks down students' engagement into three parts: behaviour, emotions and cognition; among which, the second part is concerned with whether or not a student actively participates in learning by exercising their mental energy to comprehend knowledge profoundly through self-study techniques, as well as exercising personal regulation of the entire learning process [2]. It has considerable Value for research at present; Although it classifies involvement as participation or activities respectively; a student who frequently responds to AI prompts might be highly behaved because he is not thinking critically; while his cognitive involvement could decrease; merely observing the output without any further engagement or reflection. In practice, this distinction is necessary; that is, many current indicators for measuring students' interest in using educational AI are based on behavioristics and cannot reflect the underlying cognitive levels reflected by these behaviourisms.

According to Sweller's cognitive load theory, the third theoretical basis is provided for describing in detail under what Conditions effective Instructional Design can help promote Learning [3]. The theory separates it into three types: intrinsic cognitive load - related to the inherent complexity of what needs to be learnt; Extrinsic cognitive Load - due to a poor design in instruction, causing working memory usage without helping learn; And Germane Cognitive Load - linked to schema formation and automated processes where new knowledge is incorporated into long-term storage. Adaptive generation of explanation contents based on the learner's initial state, and providing a structureised response to their questions can reduce the teacher's workload by hand-ocuping tasks; They also bear the risk of lowering intrinsic cognitive load so low as to eliminate productive difficulty leading to schema construction; thus, only performing surface engagement and not truly engaging with it at a deep level. Bransford, Brown and Cocking synthesised their own studies on the research results in learning science to reinforce this fear: Learning environments need to have both support and difficulties; That is not impossible for students to learn without encountering any problem[5].

Generally speaking, the research on the application of artificial intelligence in education focuses on three types: ai-guided learning, ai-assisted learning and ai-enabled learning; They serve as a means for analysis to place generative AI among other forms of educational technology over time and distinguish its nature compared with past iterations. Cope and Kalantzis's fundamental work on multiliteracies places the multi-modal dimension of this analysis within a larger project to understand how meaning is constructed through various representation Systems in modern learning contexts [6].

3. Generative AI and Multimodal Learning Environments

The new generative AI system's reconfiguration of the multimodal learning environment differs significantly from previous educational Technologies. Whereas traditional multimedia learning research developed its theories based on the instruction material with a fixed multimodal structure at that time point; generated-AI generates output results with dynamic changes in modality, complexity and sequence due to learner inputs. A student requesting an explanation of a complicated biological process would receive a text account along with a generated diagram; then, following this presentation, there is often some form of dialogue where the AI addresses any ambiguities raised by the students' understanding-multimodal interaction that cannot be realised through pre-coded instruction materials and relies not only on the effectiveness of the AI-generated content but also how well the student engages with it.

Directly linked to all aspects of the basic processing law system proposed by Meyer[1]. The separate-channel assumption proposes that, in terms of generating texts and images simultaneously with a unified concept target, generative AI may be less susceptible to loss from attention-splitting when verbal and visual information are not well coordinated during training. The constrained capacity assumption points out that under an AI system optimised for breadth rather than accuracy, its output could exceed a learner's handling capacity due to excessive multimodal richness, thereby causing overloads as per the coherence and redundancy constraints intended to avoid. The active processing assumption posits that learning is an activity of selecting, organising and integrating new information; therefore, the problem raised in this paper is how to ensure that the interactive dialogue style of generative AI education fosters real thinking abilities rather than merely producing an artificial sense of connection through language.

According to Zawacki-Richter et al.'s systematisation of AI applications in higher education, it is evident that although some tools can support the basic retrieval task as well as simple presentations[8], they have yet failed to facilitate the more complex tasks such as analysis and synthesis. This viewpoint is consistent with the demand for individualisation under the framework of a multi-modal learning mode; Generative AI may provide information in an intuitive manner across multiple modalities, but it cannot be assumed to prompt learners actively participate in organizing these elements into all-inclusive knowledge systems. Similarly, the analysis of Kasneci and others on the opportunities and obstacles for large language models in education points out that one of the critical issues facing educational use of generative AI is "cognitive off-loading", where learners are gradually removed from their thinking tasks through AI systems; They believe this can be an educational concern as pronounced as those associated with its integration[10].

4. Cognitive Engagement Mechanisms in AI-Mediated Learning

The mechanism by which generative AI supports deep cognitive engagement is real, but it depends on certain conditions: under these Conditions, when there is an activation of Cognitive processes within a Learning Environment that generates Understanding according to the identification criteria in both the Engagement and Learning literatures; Conversely, When AI Tools replace these Processes entirely instead of supporting them.

Among them, the promising one is that generative AI can assist in elaborative inquiry, a process where students deepen their understanding of new content by producing explanations, making links with existing knowledge, and recognizing the conditions for generalisation. When students interact with generative AI in this way, they pose questions that explore the limits of their knowledge and use AI responses to prompt further exploration; That is to say, during such an interactive process, it helps promote the active thinking that meets the

requirements of Meaningful Multimodal Learning proposed by Mayer [1]. According to Holstein et al.'s research on the design features of human-AI collaboration in student-oriented education systems, it is necessary to build an AI system with prompting function for learners' reflections, while avoiding resolution of cognitive conflicts by artificial intelligence itself [9].

Secondly, the Mechanism of Enhancing Comprehension Through Metacognitive Activation. generative-AI system prompting students to express their existing understanding before offering explanations, predicting results in advance after being prompted, evaluating the quality of AI-generated responses according to one's own growing sense of knowledge development, etc., can help build a feedback loop for self-monitoring and self-control based on cognition[2]. Self-restraint in cognitive involvement has been proven to be critical, and whether students view interactions with AI as opportunities for assisted intellectual work or as ways out without any thinking are key issues here. Ouyang and Jiao's Framework for AI-enhanced learning indicates that the pedagogical Design of interaction between humans and AI systems - The degree to which it positions learners as active participants is more crucial in terms of cognitive motivation compared with how sophisticated the AI technology is [7].

Adaptive scaffolding of cognitive-load management falls into the third kind. Accordingly, generative AI could adjust its generation difficulty level based on learners' knowledge bases and maintain a Zone of Productive Difficulty (ZoPD) by being neither too simple nor too advanced to frustrate or challenge fully independently. According to Sweller's division of internal and external cognitive load theory, as a basis for calibration; An effective AI scaffold can decrease the second one by presenting information clearly and systematically so that it remains higher after retaining complicated knowledge structures promotes schema construction[3]. There are two problems in practice: Firstly, at present's generative-AI-systems have no ability to distinguish between creative and summary tasks; Students will replace part of their work with AI-generated content entirely when using it - generating an entire answer instead of only a partially scaffolded one or reading just a summary without understanding the main points.

5. Discussion

Through this paper's development, it is not uncritically enthusiastic nor do we reflexively doubt any aspect of the connection among generative AI and cognitive engagement. Through a multimodal Learning Framework, it reveals real Mechanisms By Which Generative AI Can Help Support Deep Cognitive Processing; And Through an Engagement Literature, Conditions Are Found Under Which Those Mechanisms May Operate. However, the same frameworks can also reveal issues with AI-assisted learning environments that focus on accessibility rather than productive difficulties or behavioural engagement lacking in-depth cognition.

The main risk is known as the fluency illusion; i.e., when we interact with artificial intelligence-generated content in a state of high fluency and comfort that makes us feel certain but lack genuine understanding. According to Bransford, Brown and Cocking's summary of the theory of learning, if a learner finds their work "easily doable", then there is no incentive for them to continue exploring more challenging content; however, it has been proven by subsequent studies that such difficulties in understanding actually enhance learners' cognitive processing capabilities[5]. Optimised generative AI systems, which aim at enhancing users' experiences, by generating responses that feel supportive, understandable, and satisfactory, may have a negative effect on the productive struggles through subtle but profound changes in how these artificial intelligence system interactions manifest as intuitive sense of comprehension.

Equity Dimensions for Generative-AI Adoption in Educational Settings Should Be Explicitly Addressed. According to Zawacki-Richter and colleagues' review, there is considerable inconsistency in the availability of artificial intelligence (AI) tools and digital competence among different student groups; therefore, it can be inferred that the cognitive-engagement enhancement brought about by good designed AI-learning-environment may be more favourable for students with higher levels of autonomous regulation and technological proficiency. Students without the necessary prerequisite knowledge will more likely use generative AI tools in an unhelpful way by replacing their own problem-solving efforts with generated material, which only reinforces the students' initial learning deficits through a superficial approach. Based on the aforementioned shortcomings, how can we guarantee that all students have equal access to AI technologies while helping teachers improve their own teaching skills and facilitating a wider spread of AI applications?

The function of the teacher in AI-mediated learning Environment Should be considered Based on this research. Kasneci et al. believe that teachers' roles change in generative AI environments; They were once merely primary knowledge suppliers but now become designers of learning scenarios - That is, they need to create tasks, dialogues and examination methods that impact whether AI aids or replaces cognition.[10] The re-conception necessitates the creation of a novel form of teaching ability beyond what most teachers' education programmes can offer; thus, establishing professional development systems that nurture these capacities should be prioritised by schools aiming to merge generative artificial intelligence with integrated creativity.

6. Conclusion

Generative AI has both positive effects and some degree of influence on students' cognitive participation; its impact varies under different conditions based on the education Environment within that context to meet the requirements for cognitivE. Developing from a multi-modal learning perspective, this work explores how generative AI has promoted deeper understandings of something; in addition, it raises issues currently faced during the practical implementation of such an Environment.

The theoretical synthesis presented here indicates that realising the educational potential of generative AI requires paying attention to the following three mutually reinforcing factors simultaneously: Educational design uses AI's multi-modal capabilities to reduce extraneous cognitive load, but preserves the core difficulty that motivates learners; Pedagogical framework makes students become active participants in AI interactions instead of passive consumers of AI-generated materials; The way for evaluating recognises whether a student is genuinely engaging through their cognitive participation versus mere behavioristic response produced by fluency in using AI.

Future Research Should Carry Out Longitudinal Empirical Studies on Cognitive Engagement In Generative AI Learning

Environments Across Diverse Subject Domains And Student Populations To Build Measurement Frameworks For Distinguishing Surface From Deep Engagements Under AI-Mediated Situations. Here, the theoretical foundations have provided some backing for the empirical study; however, translating these theoretical insights into practice-oriented pedagogical recommendations still needs more ecological-valid research over an extended period than has been conducted so far in this domain.

References

[1] Mayer RE. Multimedia Learning. Cambridge University Press; 2001.

[2] Fredricks JA, Blumenfeld PC, Paris AH. School engagement: Potential of the concept, state of the evidence. Review of Educational Research. 2004;74(1):59–109.

[3] Sweller J. Cognitive load during problem solving: Effects on learning. Cognitive Science. 1988;12(2):257–285.

[4] OpenAI. GPT-4 Technical Report. OpenAI; 2023. Available from: https://arxiv.org/abs/2303.08774

[5] Bransford JD, Brown AL, Cocking RR, editors. How People Learn: Brain, Mind, Experience, and School. National Academies Press; 2000.

[6] Cope B, Kalantzis M. A pedagogy of multiliteracies: Designing social futures. Harvard Educational Review. 1996;66(1):60–92.

[7] Ouyang F, Jiao P. Artificial intelligence in education: The three paradigms. Computers and Education: Artificial Intelligence. 2021;2:100020.

[8] Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education. 2019;16(1):39.

[9] Holstein K, McLaren BM, Aleven V. Improving the effectiveness of human-AI collaboration in student-facing educational AI. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. ACM; 2019.

[10] Kasneci E, Seßler K, Küchemann S, et al. ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences. 2023;103:102274.