Beyond the "Wow" factor: Using Generative AI for Increasing Generative Sense-Making

Guido Makransky
gm@psy.ku.dk

University of Copenhagen https://orcid.org/0000-0003-1862-7824
Ban M. Shiwalia
University of Copenhagen https://orcid.org/0000-0001-8057-5303
Tue Herlau
Danish Technical University https://orcid.org/0000-0001-7288-6953
Steven Blurton
University of Copenhagen https://orcid.org/0000-0003-4311-0202

Research Article

Keywords: generative artificial intelligence, generative sense making, learning, teaching, self-explaining

Posted Date: December 12th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-5622133/v1

Additional Declarations: The authors declare potential competing interests as follows: Tue Herlau is a Researcher at The Danish Technical University and Founder of ChatTutor

Abstract

Generative Artificial Intelligence (GenAI) has emerged as a transformative tool in education, offering scalable, individualized learning experiences. However, there is a notable lack of theoretically informed and methodologically rigorous research on how GenAI can effectively augment learning. This study addresses this gap by investigating the potential of a theoretically informed GenAI chatbot, ChatTutor, to facilitate generative sense-making, leveraging principles from generative learning theory. The study had two primary goals: first, to build on theory to propose how GenAI could be used to support generative sense-making; and second, to empirically test its impact on conceptual knowledge, self-efficacy, and trust immediately after the intervention, as well as on conceptual knowledge, enjoyment, and behavioral intentions in a follow-up test four weeks later. Conducted in an authentic university course, the pre-registered experiment with 175 students compared ChatTutor to a generic GenAI system (ChatGPT) and a teaching-as-usual condition. Results show that ChatTutor significantly enhanced trust, enjoyment, and behavioral intentions but not self-efficacy. While ChatTutor improved conceptual knowledge over ChatGPT immediately after the intervention, it did not outperform teaching-as-usual. Four weeks later, ChatTutor significantly outperformed teaching-as-usual but not ChatGPT in conceptual knowledge. The manuscript underscores the importance of integrating human-centered design and educational psychology theories into GenAI applications to optimize learning outcomes and proposes future research and practical implications.

Introduction

Generative Artificial Intelligence (GenAI) has received increasing public interest and research (Kasneci et al., 2023; Lorenz et al., 2023; Yan et al., 2024). The widespread use of GenAI has surged with the release of publicly available large language models such as BERT (Devlin et al., 2019), Llama (Touvron et al., 2023), OpenAI (Helmore, 2023), Claude (Anthopic, 2024), and GPT-4 (Edwards, 2023). These models generate text, images, audio, and structured synthetic data through user-friendly interfaces, while also producing human-like responses (Cuéllar et al., 2024; Schramowski et al., 2022). Compared to predictive AI, which can be used for predicting labels or values, GenAI is designed to create new content or ideas and express them in real-time conversations (Qadir, 2023). GenAI has taken society by storm, in the history of the internet, no consumer app has grown faster than OpenAI's ChatGPT (Chow, 2023). The experience of asking a Chatbot anything and receiving a coherent, face valid answer has been described by users as "mind-blowing", "impressive", and "amazing" (Taecharungroj, 2023). This has led many to anticipate that this technology will inevitably revolutionize society, reshaping the ways we work, live, and learn (Lorenz et al., 2023). In education this could change how teachers teach, how learners learn and are assessed, and how education is administered (Chiu et al., 2023). However, emerging technologies typically progress through different levels of hype, and initial expectations are often exaggerated due to media coverage (Fenn & Blosch, 2020). While it can be argued that some emerging technologies such as computers have eventually had a significant influence on education, these changes are usually slower than initially expected and may be detrimental to the learning process (Chandler, 2009). In general, the digitalization of society results in challenges and opportunities for learning and education (Fischer et al., 2020; Yan et al., 2024).

A common challenge with the introduction of novel technology is that it is often implemented and examined from a technocentric perspective, without incorporating extensive research and theory from educational psychology and the science of learning (Chandler 2009; Brennan, 2015; Yan et al., 2024). In this study, we argue that it is important to look beyond the "wow" factor of GenAI in education and adopt a learner-centered approach to investigating how the technology can enhance rather than distract from learning. Here, we agree with the perspective that psychological theory is central to understanding how GenAI can be used in this endeavor (e.g. Molenaar, 2022; Yan et al., 2024). Specifically, we use generative learning theory (Fiorella, 2023; Fiorella & Mayer, 2016) to design a GenAI educational application that supports human-AI collaboration in promoting generative processing. Generative learning activities (GLAs) are learner-driven actions that foster the active construction of knowledge as they prime key cognitive processes such as selecting, organizing, and integrating knowledge (Fiorella & Mayer, 2015, 2022; Mayer, 1996; Wittrock & Farley, 2010). By aligning GLAs with the capabilities of GenAI, we explore how integrating human and machine intelligence can enhance learning outcomes, augmenting rather than replacing human abilities (Molenaar, 2022). That is, we adopt an augmentation perspective, viewing AI not as a simulation of human intelligence but as a distinct form of intelligence that complements our own, serving as a tool to potentially enhance human capabilities and help us function better as humans (Molenaar, 2022). Combining human and AI capabilities through hybrid intelligence has been proposed as a key to advancing the learning sciences in an AI era (Järvelä et al., 2023). But systematic research that builds on examining this integration is scarce (Hennessy et al., 2024; Roschelle et al., 2020). Ideally, hybrid intelligence systems should optimize human strengths while compensating for their weaknesses.

In a world where knowledge is abundant and easily accessible, the focus of education could shift from imparting knowledge to nurturing students' ability to ask meaningful questions and critically evaluate information. A meta-analysis on the effect of inducing self-explanation found that having learners generate an explanation is often more effective than presenting them with an explanation (Birsa et al., 2018). Its impact was greatest when learners used new information to revise their explanations. However, while instructor-scripted self-explanation prompts work well for standardized content, they are impractical for highly personalized learning, and the authors recommend investigating computer-generated, content-specific prompts. GenAI makes this possible at scale, and in this study we investigate its potential to foster higher-order cognitive processes, such as conceptual understanding (Anderson & Krathwohl, 2001), while also promoting self-efficacy (Bandura, 1977), trust (Gulati et al., 2019), enjoyment (Davis et al., 1992; Huang & Zou, 2024), and behavioral intentions (Lee & Choi, 2017).

Methodological rigor has been recently highlighted as a serious challenge to the research investigating GenAI in education (Yan et al., 2024). A criticism is that the majority of research on GenAI in education is conceptual, focusing on literature reviews, analyses of public discourse, or surveys of people's perceptions (Hennessy et al., 2024, p. 6). Emerging evidence also suggests that the impact on learning and engagement presents a complex picture, such as instances where GenAI may have a positive initial impact on learning outcomes, however, it can act as a crutch and the removal of GenAI may ultimately hinder learning (Darvishi et al., 2024; Nie et al., 2024). This has prompted calls for more experimental research in authentic educational settings, grounded in robust theoretical frameworks, as well as for longitudinal studies to evaluate GenAI's long-term benefits to human learning by comparing its effectiveness with conventional methods (e.g. Hennessy et al., 2024; Yan et al., 2024). Furthermore, for GenAI to enhance rather than detract from learning, collaboration among researchers, practitioners, GenAI developers, policymakers, and educators is crucial to ensure its effective and responsible integration into teaching, aligning with educational goals (Yan et al., 2024).

The first goal of this paper is to use theories from educational psychology and technology-enhanced learning to provide an understanding of how GenAI can effectively be incorporated into a lesson. While many theories could be used as an approach for using GenAI in education, in this article we focus on generative learning theory (Fiorella, 2023; Fiorella & Mayer, 2016). The second goal of the paper is to empirically test if the use of GenAI in education implemented to facilitate generative learning can increase conceptual knowledge, self-efficacy and trust immediately after a lesson, as well as conceptual knowledge, enjoyment and behavior intentions in a follow-up test four weeks later. In an experiment conducted in an authentic university setting, we compare the use of a theoretically informed GenAI application with a standard GenAI application and a teaching-as-usual condition. Prior to describing this experiment in more detail, we begin with the first objective and provide a theoretical overview of how GenAI can be used to facilitate generative learning.

Theoretical Background

Generative Learning

Generative learning involves "making sense" of learning material by actively organizing and integrating it with existing knowledge (Wittrock, 1989). According to Fiorella (2023, p. 2) "*a generative learning activity (GLA) is something a learner does to try to make sense of what they are learning*". There is evidence that learners do not spontaneously engage in sense-making when learning from text (e.g. Fiorella & Mayer, 2017), visualizations (e.g. Makransky & Petersen, 2021) or examples (Renkl, 1997). Prompting learners to engage in GLAs is therefore a potential remedy for achieving meaningful and durable learning outcomes (Fiorella, 2023). Fiorella and Mayer (2015, 2016) identified eight generative learning strategies (GLS) including: summarizing, mapping, drawing, imagining, self-testing, self-explaining, teaching, and enacting. Reviews and meta-analyses generally show that GLS support learning but each is susceptible to potential boundary conditions related to learner characteristics, learning materials, and support levels (Bisra et al., 2018; Brod, 2021; Dargue et al., 2019; Fiorella & Zhang, 2018; Lachner et al., 2021). That is, some learners need extensive guidance (Fiorella & Zhang, 2018), and in this study we investigate if GenAI can be used to provide personalized guidance that can improve educational outcomes.

In the current study we are specifically interested in the verbalizing generative learning activities of teaching which incorporates elements of self-explaining. Generating explanations during learning can activate prior knowledge, integrate new information, enhance memory, and encourage reasoning by prompting students to make inferences and revise mental models (Brod, 2021). A recent meta-analysis (Bisra et al., 2018) found a mean effect size of 0.55 SDs for self-explanation, comparing conditions with and without self-explanation prompts. When considering only studies in which time on task was equated, the effect size was still 0.41 SDs. Generating explanations during learning can be done through self-explanation, which involves generating verbal statements (involving inferences) to clarify the meaning of the learning material to oneself (Fiorella & Mayer, 2022, p. 341). Fiorella and Mayer's (2015) review found positive effects of self-explaining in 44 of 54 experiments, with a median effect size of 0.61.

Self-explaining through teaching involves generating verbal statements (involving inferences) to convey the meaning of the learning material to others (Fiorella & Mayer, 2022, p. 341). Fiorella and Mayer's (2015) review reported positive effects of learning by teaching in 17 of 19 experiments, with a median effect size of $d = .71$. A meta-analysis by Kobayashi (2019) found stronger effects for students who actually taught ($g = .56$) compared to those who only prepared to teach ($g = .35$). Explaining to others adds social elements, including a sense of presence and opportunities for interaction, for instance by answering questions.

Evidence suggests that explaining is most effective when done aloud, without access to learning materials, when students possess background knowledge, and when they respond to thought-provoking questions (Fiorella & Mayer, 2022). Supporting factors include scaffolded peer interactions (King, Staffieri, & Adelgais, 1998), focused prompts (Berthold, Eysink, & Renkl, 2009), and explicit training (McNamara, 2004). However, most of this research has been conducted without the ability to provide adaptive specific feedback that is now possible with GenAI. In this study, we investigate the GSL of teaching, building on broader research in self-explanation, generative learning, AI-based education (AIEd), among other related fields.

How can the use of Theoretically Informed GenAI Increase Generative Sense-Making?

It is possible for instructors to use GLA products as formative assessment and adapt instruction and guidance as necessary (van de Pol et al., 2020). For instance, when learners' GLA products show knowledge gaps or misconceptions, instructors can give targeted feedback to aid their internal sense making, improve the GLA products, and enhance learning outcomes (Fiorella, 2023). The challenge is that assessing knowledge gaps and misconceptions and then providing targeted feedback is time consuming and difficult in most educational settings. The question that we attempt to investigate in this study is if a GenAI Chatbot can be used as a tool to scaffold learning by providing personalized feedback to stimulate and support generative learning activities.

According to a taxonomy developed by Holmes and Tuomi (2022), AIEd applications can be classified into those that are student-focused, teacher-focused, or institution-focused. This study exclusively deals with student-focused AIEd applications that are leveraged to support student learning. Holmes & Tuomi (2022) also distinguish between AI-tools designed *for* students and those that are used *by* students. In this study we compare a theoretically informed GenAI-tool designed for students (ChatTutor) and a GenAI-tool used by students (ChatGPT) to a teaching-as-usual condition. ChatTutor builds on generative learning theory, and technology enhanced learning evidence and theory (e.g., Fiorella, 2023; Fiorella & Mayer, 2022; Mayer, 2014). Figure 1 depicts a framework for the development of a student-focused GenAI educational application where interaction is initiated by a) the learner, b) the AIEd system, and c) the instructor. In this study both ChatTutor and ChatGPT allow for interaction initiated by the learner. However, building on the research from generative learning (Fiorella, 2023; Fiorella & Mayer, 2016), we have furthermore designed ChatTutor to prompt students to engage in GLA and provide formative feedback. Effective guidance involves adapting learning materials and GLA support to learner's knowledge and beliefs, thereby enhancing their internal sense making and external products (Fiorella, 2023). We suggest that GenAI based learning systems such as ChatTutor can use the quality of these products as formative assessments to inform further adjustments of learning materials and GLA support. This is the case because in addition to providing responses to prompts that are initiated by the learner, learners who do not automatically initiate sense-making can systematically be prompted to do so by the GenAI system. We will use existing literature related to GLS to propose how this can be done in the following sections.

Using a Theoretically Informed GenAI Chatbot for Prompting and Guiding GLA.

There is evidence that for novices, complete discovery without guidance is ineffective; and some level of scaffolding is necessary (Mayer, 2004; Pedaste et al., 2015). Therefor for student-centered learning to be effective, guided discovery is essential. Ideally, this guidance must be personalized, which can be facilitated through technological advances (De Jong, 2006). GenAI systems can provide valuable feedback, and the focus should include the knowledge students possess as well as the questions they ask (Berlyne, 2006; Dillenbourg, 1999). The challenge is to support students in asking the right questions, finding accurate answers, and critically analyzing the information they receive to further investigate and understand the knowledge (Hmelo-Silver et al., 2007). Many guided learning principles including problem-based learning and inquiry learning, extensively use scaffolding to reduce cognitive load, enabling students to learn in complex domains (Hmelo-Silver et al., 2007). Fiorella (2023) suggests that GLA support can be categorized into three broad levels: explicit instruction, scaffolded practice, and independent practice which align with research on GLAs (Fiorella & Zhang, 2018; McNamara, 2017) and cognitive load theory (Sweller et al., 2019), progressing from worked examples to partial problems to problem-solving practice.

Explicit instruction includes explanations and demonstrations on selecting and performing GLAs. Explicit prompts help learners to generate internal self-explanations (Fiorella, 2023). The literature on worked examples suggests that this is necessary because learners often struggle to detect errors (Van Meter, 2001). They may need support in generating internal feedback (e.g., self-explaining why their response was inaccurate) and revising their knowledge (Zhang & Fiorella, 2023). Wittwer and Renkl (2010) describe how effective instruction involves a balanced combination of providing support, such as worked solution steps, and encouraging learner activity, like self-explanation prompts. An experienced instructor can provide this balance by encouraging learners to engage in a GLA such as self-explaining, or teaching, and then help them detect gaps in their knowledge thereby stimulating further inquiry. This requires the ability to engage learners in generative processing, and the knowledge to recognize and highlight gaps in a way that motivates students to engage in further information seeking (Murayama, 2022). A GenAI chatbot can be programmed to provide specific self-explanation prompts, trained using a targeted knowledge bank, and interact with learners as a knowledgeable teacher, an overconfident peer, or a curious collaborator.

Once learners become more familiar with the course content then GenAI based chatbots can progress to scaffolded practice activities. Scaffolded practice includes providing specific prompts or hints (Berthold & Renkl, 2009; Roelle & Berthold, 2017) or partial representations for learners to complete (Fiorella et al., 2021; Ponce et al., 2020; Schwamborn et al., 2010). Learners can benefit from scaffolded prompts where they fill in key components of an explanation rather than generating the entire explanation themselves (Bai et al., 2022). Self-explaining is more effective when learners receive focused prompts targeting specific relationships or principles rather than open or generic prompts (Berthold & Renkl, 2009). Scaffolded practice is most effective when followed by appropriate feedback, which helps learners evaluate and correct their performance (Johnson & Marrafano, 2022). In the current study we build on this literature by using GenAI to prompt students to engage in the generative activities of teaching. More specifically, students are given prompts that ask them to explain key components or explanations. The ChatTutor chatbot engages students in personalized written conversations based on the course slides and texts, and responds to address the specific answers provided by students and correcting any misconceptions or gaps in knowledge. Students are gradually provided with more information, scaffolding their understanding until they can offer a suitable explanation. This approach simulates a student-tutor interaction, where the tutor encourages generative learning by prompting students to explain specific content, intervening only when misconceptions are evident or when students request help.

Finally, independent practice involves repeated opportunities to use GLAs in specific learning contexts. This practice is crucial for learners to use GLAs spontaneously (Manalo et al., 2017) but is not investigated in this study.

It is important to note that we are not proposing that GenAI is necessary or in its current technological iteration optimal for progressing from explicit instruction through scaffolded practice and ultimately to independent practice, however, we do propose that GenAI can be used to scale this process when resources to formatively assess students and provide personalized support is not possible. In the following section we describe how a theoretically informed GenAI chatbot can be used to mitigate some of the documented barriers to GLS.

How can a Theoretically Informed GenAI Chatbot Mitigate Known Barriers of GLS?

Fiorella (2023) describes different cognitive, metacognitive and motivational barriers to sense making. Below we describe how a GenAI chatbot could attempt to mitigate these barriers thereby increasing sense-making and ultimately learning and motivational outcomes.

Cognitive Barriers: One cognitive barrier to engaging in effective GLS is insufficient background knowledge or available memory capacity to engage in GLAs effectively (Castro-Alonso et al., 2021). Learners need sufficient background knowledge to generate the necessary inferences for sense-making (Simonsmeier et al., 2022; Willingham, 2008). Without it, learning materials won't be effective, regardless of the GLA used. With GenAI, students can ask a personalized chatbot particular questions thereby gaining specific knowledge. In ChatTutor this is implemented by linking directly to the texts and slides from the course., or by asking the LLM to provide simpler alternative explanations such as summaries or explanations. This is relevant because lower-knowledge students often benefit more from GLAs than higher-knowledge students (Fiorella & Mayer, 2015). This is because higher-knowledge learners are more likely to engage in sense-making spontaneously (Lombrozo, 2006).

Metacognitive Barriers: Metacognitive barriers include a lack of strategic knowledge about what GLAs are, how to choose the right ones, and how to use them effectively. Many learners do not use GLAs spontaneously when studying texts (Fiorella & Mayer, 2017), visualizations (Renkl & Scheiter, 2017), worked examples (Chi et al., 1989; Renkl, 1997), solving problems (Rellensmann et al., 2022), discussing material with peers (Roscoe & Chi, 2007), or writing essays (Bereiter & Scardamalia, 2013). GenAI could provide explicit instructional support to assist learners to engage in GLAs. Specific prompting and scaffolding have been shown to be more effective compared to generic prompts in explaining, visualizing, or enacting concepts (Berthold & Renkl, 2009; Schmeck et al., 2014). Moreover, the quality of what learners generate during learning typically predicts their performance on later comprehension and transfer tests (Chi et al., 1989; Fiorella & Kuhlmann, 2020; Goldin-Meadow et al., 2009; Renkl, 1997; Schwamborn et al., 2010).

Motivational barriers: Motivational barriers include beliefs about one’s ability to use GLAs successfully, the perceived value of GLAs for achieving one’s goals, or the perceived cost of GLAs (e.g., (Schukajlow et al., 2022). GenAI could increase self-efficacy by providing individualized mastery experiences (Bandura, 2001). A GenAI Chatbot can be seen as a more knowledgeable companion, providing positive feedback and enhancing learning by targeting the learner's zone of proximal development (Vygotsky, 1978).

Can a Theoretically Informed GenAI Chatbot influence other outcomes?

In addition to cognitive learning outcomes and self-efficacy, there are other important outcome variables that can ultimately impact the use of GenAI for learning. In this study we selected three: trust, enjoyment, and behavioral intentions.

The adoption of GenAI in educational contexts brings forth several ethical challenges, such as transparency, privacy, equality, and beneficence (Khosravi et al., 2022). While all of these factors are relevant, building a GenAI chatbot that takes a learner-centered approach to interaction and builds on the curriculum in a course may specifically be relevant for transparency. A recent systematic review by Yan et al. (2023) found that most GenAI tools (92%) currently employed to support learning are comprehensible only to AI experts. Educators, students, and other key stakeholders often lack the necessary insight into how these tools function. The authors highlight that the transparency gap largely stems from the limited integration of human-in-the-loop approaches in prior research. For instance, educators and students have seldom been actively involved in the design and evaluation of GenAI-based educational technologies. This shortfall highlights the urgent need for learner-centered AI, emphasizing the importance of engaging all stakeholders in the development process to ensure GenAI tools are both effective and meaningful in real-world educational settings. In this study we engage instructors, researchers, learners, and a GenAI development start-up in developing the ChatTutor system, and assess students’ trust (Gulati et al., 2019) in using the system in our experiment.

It can also be useful to measure the enjoyment of using AI technology in education because enjoyment is a powerful motivational factor that directly impacts learners' engagement and long-term adoption of these tools (Huang & Zou, 2024). According to the Control-Value Theory of Achievement Emotions, enjoyment plays a critical role in fostering intrinsic motivation and engagement by enhancing learners' perceptions of control over their learning and the value they assign to educational tasks (Pekrun, 2006). This positive emotional experience can reinforce satisfaction with the learning process and increase the intention to continue using AI-enhanced educational platforms (Venkatesh & Bala, 2008), leading to sustained engagement and ultimately more widespread adoption (Dai et al., 2024). We therefore measure trust, enjoyment, and behavioral intentions in this study.

Current study

In the current study we address some of the gaps in the literature by designing a GenAI application building on a learner-centered theory of learning and instruction. Our main research question (RQ 1) is: What are the effects of using a theoretically informed GenAI chatbot (ChatTutor) to stimulate generative sense making on learning outcomes? We pre-registered two main hypotheses regarding this research question:

- **Hypothesis 1** (H1): Learners in the theoretically informed GenAI condition (ChatTutor) will achieve higher scores on a conceptual knowledge test immediately after the lesson when compared to learners in the teaching-as-usual (H1a) and ChatGPT-4 (H1b) conditions.

- **Hypothesis 2** (H2): Learners in the theoretically informed GenAI condition (ChatTutor) will achieve higher scores on a conceptual knowledge test administered four weeks after the lesson when compared to learners in the teaching-as-usual (H2a) and ChatGPT-4 (H2b) conditions.

Our secondary research question (RQ 2) is: Does the use of a theoretically informed GenAI chatbot (ChatTutor) lead to higher (1) self-efficacy compared to the teaching-as-usual condition and the ChatGPT condition, and more (2) trust, (3) enjoyment, and (4) intentions to use the technology in the future compared to the ChatGPT condition?

Methods

Study design

The study adopted a between-subject design consisting of one control group and two distinct GenAI conditions: ChatGPT and ChatTutor. The design, hypotheses, data collection, and analysis plan for this study were pre-registered on AsPredicted in September 2024, prior to the commencement of data collection. The anonymized preregistration is available at: https://aspredicted.org/w5mk-4vx4.pdf. The study was conducted at a large European University. Based on the self-assessment, the department's local ethical board accepted the research project.

Procedure

The project was initiated by psychology university instructors who were interested in investigating if GenAI could improve conceptual knowledge retention, self-efficacy, trust, enjoyment, and behavioral intentions in a cognitive psychology university course on the topic of Sternberg's serial memory scanning experiment (Sternberg, 1966, 1969) which is a challenging topic for most students. The cognitive psychology course consists of a weekly lecture with over 240 psychology students, followed by both a seminar and an exercise class. At the beginning of the semester, students were randomly assigned to one of seven cognitive psychology classes with approximately 35 students in each. In accordance with local study board regulations, attendance in the lecture or the classes was not mandatory.

The experiment took place in the first week of October, 2024, when the topic in the course was short-term memory and working memory. The follow-up took place as the first part of the lesson four weeks later. The Sternberg task was first introduced at the end of the seminar classes. In the following exercise class, students received a more in-depth overview and performed the task themselves. The exercise class consisted of three 45 minute blocks with 15 minute breaks between. The first block started with a wrap-up of the previous week, followed by a more in-depth introduction into the short-term memory task. In the second teaching block students took part in the Sternberg experiment themselves. After running the experiment, the students analyzed their logfile created at the end of the experiment, this task was otherwise unrelated to the current study. The third block started with a 20-minute power-point presentation by the instructor which was standardized across the different classes. This was followed by a 15 minute follow-up activity which was the focus of the experiment. In this between-subjects pre-registered experiment, two exercise classes were assigned to using the theoretically informed GenAI system (ChatTutor), two classes were assigned to the standard ChatGPT-4 (standard GenAI) condition which had the same interface as the ChatTutor condition, and the remaining three classes were assigned to the teaching-as-usual condition. Therefore, the difference between the groups was isolated to the generative learning activity which took approximately 15 minutes (see Fig. 2). The final 10 minutes of the class consisted of a post-test including measures of conceptual knowledge, self-efficacy, and trust in AI. Four weeks after the intervention students used 10 minutes at the beginning of the exercise class to take a follow-up test that included another conceptual knowledge test as well as measures of enjoyment and behavioral intentions. Students were informed by their instructors that the assessment was an effort to investigate how to improve the teaching methods used in the course, but the students were not informed of the differences between instructor classes. At the end of the semester, students were informed of the results of the experiment and students in all treatment conditions were given access to the ChatTutor system.

Participants

A total of 175 university students who were enrolled in the cognitive psychology course and attended their exercise class, agreed to participate in the study. Of these, 75 participants were in the teaching-as-usual condition, 49 were in the ChatGPT condition, and 51 were in the ChatTutor condition. The topic of the lesson was centered on the Sternberg Experiment. The experiment was integrated into classes' schedules and conducted during regular teaching hours. Participants provided their informed consent prior to the commencement of the experiment. As outlined in the preregistration, data from participants who chose not to engage in the experiment (< 5) were excluded. To our knowledge there were no additional students who opted not to use the AI program, and we did not exclude any students due to technical issues while using the GenAI program. A total of 124 students who completed the initial test, also completed the delayed follow-up test four weeks after the intervention. Among these, 54 participants belonged to the teaching-as-usual condition, 32 to the ChatGPT condition, and 38 to the ChatTutor condition.

Conditions

The experiment was isolated to the follow up activity which lasted approximately 15 minutes. All students received the similar instructions regarding the follow-up activity. Those in the teaching-as-usual-condition had 15 minutes to reflect over what they had learned and could discuss this with their classmates and had the opportunity to ask their instructor any questions which they were in doubt about regarding the topics that they had learned that day. The students in the ChatGPT condition had 15 minutes to reflect over what they had learned and had the opportunity to ask the GenAI any questions which they were in doubt of regarding the topics that they had learned that day. The students in the ChatTutor condition also had 15 minutes to engage in the ChatTutor system as described above.

1. Teaching-as-usual: Students in this condition were able to engage with the material and discuss with their fellow-students and could ask the instructor any relevant questions related to the material. Moreover, they were free to use any existing learning materials and internet sources.
2. Generative Sense Making (ChatTutor) condition: Students in this condition got access to ChatTutor website, which was designed to encourage generative processing as their learning tool. Students were encouraged to engage with the tool which actively prompted students to partake in the generative learning activity of teaching with scaffolded feedback. The ChatTutor system allowed learners to escalate a question to a human instructor if they encountered confusing or inaccurate information.
3. Standard GenAI (ChatGPT) condition: Students in this condition got access to a website interface that was identical to ChatTutor. However, the system linked directly to ChatGPT-4. Students were encouraged to engage with the tool with the same general instructions as in ChatTutor condition but the system responded based on the current technological capabilities of ChatGPT-4. However, the system also allowed learners to escalate a question to a human instructor if they encountered confusing or inaccurate information.

Measures

The complete list of items is included in the supplementary materials.

Conceptual Knowledge (post-test and follow-up): A test assessing participants' conceptual knowledge, with questions related to the memory scanning experiment and Sternberg's original hypotheses (Sternberg, 1969) was administered immediately after the intervention/at the end of the lesson and again after four weeks. The test consisted of 10 multiple choice items, however, one was removed from the analyses and not included in the follow-up test because it accidentally contained more than one correct answer, resulting in a total of nine items. The test was designed to measure a broad range of topics covered with a focus on comprehension and not factual knowledge. That is, we prioritized content validity by having a broad range of questions rather than focusing on a narrow range of content which could have potentially increased internal reliability. The reliability coefficient for the nine-item conceptual knowledge test was $\alpha = .55$ in the post-test and $\alpha = .68$ in the follow-up. The test-retest reliability was $r = .41$.

New Conceptual Knowledge items (follow-up): In addition to the nine items, four new conceptual knowledge items were administered in the follow-up to ensure that we could measure conceptual knowledge while accounting for testing effect (that students would remember the answers to the items because they had seen them before). The reliability coefficient for the new conceptual knowledge items was $\alpha = .40$.

Self-efficacy (post-test): We measured self-efficacy using three items adapted from (Pintrich et al., 1991): 1. "I'm confident I can understand the concepts of serial and parallel memory search"; 2. "I'm confident I can understand the concepts of serial exhaustive search and serial self-terminating search"; 3. "I'm confident I could explain the Sternberg memory search graphs to a friend". The reliability coefficient of the self-efficacy measure was $\alpha = .82$.

Trust (post-test): We measured trust using four items adapted from Gulati et al. (2019): 1. "I feel I must be cautious when using ChatTutor"; 2. "I believe that ChatTutor will act in my best interest"; 3: "I think that ChatTutor is competent and effective in providing teaching assistance"; 4. "I can trust the information presented to me by ChatTutor". The reliability coefficient of the trust measure was $\alpha = .64$.

Enjoyment (follow-up): We measured enjoyment using three items adapted from Huang & Zou (2024), originally from (Davis et al., 1992): 1) "I find it satisfying to use ChatTutor," 2) "I find it rewarding to use ChatTutor," and 3) "I find it pleasant to use ChatTutor." The reliability coefficient for the enjoyment measure was $\alpha = .80$.

Behavioral intentions (follow-up): We measured behavioral intentions with three items adapted from Lee and Choi (2017), originally from Davis et al. (1992). 1. “I intend to use ChatTutor again”; 2. “In the future, I will use ChatTutor to support my learning processes”; 3. “I would recommend ChatTutor to others”. The reliability coefficient for behavioral intentions was $\alpha = .55$.

All items in the self-efficacy, trust, enjoyment, and behavioral intentions were measured on a 5-point scale (Strongly disagree:1. Disagree: 2. Neither agree nor disagree: 3. Agree: 4. Strongly agree: 5).

Data analysis and deviations from pre-registration

The analyses for the study were performed using SPSS Statistics (IBM Corp., 2023). We examined our hypotheses using independent samples t-tests rather than ANOVA as the hypotheses were framed as pairwise comparisons and because trust, enjoyment and behavioral intentions were only assessed in the two GenAI groups. A sensitivity power analysis using G*power (Faul et al., 2007) indicated that with an $\alpha = .05$, and power of .80 we could expect to find an effect size $d = .45$ for the difference between the ChatTutor and teaching-as-usual conditions and an effect size of $d = .50$ between the ChatTutor and ChatGPT conditions. This value increased to $d = .53$, and $d = .60$ respectively in the follow-up due to drop-out. Values below these levels should therefore be interpreted with caution. Please note that we only pre-registered our measures of conceptual knowledge, self-efficacy, and trust, but did not pre-register the measures of enjoyment or behavioral intentions, which were initially left out of the post-test due to time constraints but we were later able to include them in the follow-up.

Results

The main results of the study can be seen in Table 1. The results are organized based on the research questions below.

Table 1: Means, standard deviations, sample sizes and significance levels for the outcome variables used in the study.

Results Related to the Main Research Question (RQ 1): The main research question in this study investigated the effects of using a theoretically informed GenAI chatbot (ChatTutor) to stimulate generative sense making on learning outcomes immediately after the intervention and in a delayed follow-up test four weeks after the intervention. The top row of Table 1 illustrates that the results partially support Hypothesis 1. More specifically, the theoretically informed GenAI ChatTutor condition resulted in significantly higher conceptual knowledge compared to the ChatGPT condition in the immediate post-test. However, while the ChatTutor condition mean score was higher than the teaching-as-usual condition, this difference was not statistically significant in the immediate post-test.

The next row of Table 1 illustrates that the results also partially support Hypothesis 2.

The theoretically informed GenAI ChatTutor condition resulted in significantly higher conceptual knowledge compared to the teaching-as-usual condition in the follow-up test four weeks after the intervention. However, while the ChatTutor condition mean score was higher than the standard ChatGPT condition, this difference was not statistically significant in the follow-up test. Because it was not possible to limit the students from using the ChatTutor interface between the lesson and the follow-up test, we asked students to report on a five-point scale how often they had used ChatTutor system since the lesson (1. Not at all; 2. Seldom; 3. Sometimes; 4. Often; 5. Very often). All of the students in the ChatGPT condition reported not using it at all, and all but six students in the ChatTutor condition reported not using it at all (four reported using it seldomly, and two reported using it sometimes). We re-ran all the analyses for the follow-up test, discarding these students as a robustness check, and found that this did not change any of the findings.

Results Related to the Secondary Research Question (RQ 2): The second research question in this study was if a theoretically informed GenAI chatbot (ChatTutor) leads to higher (1) self-efficacy compared to the teaching-as-usual condition and the ChatGPT condition, and more (2) trust, (3) enjoyment, and (4) intentions to use the technology in the future compared to the ChatGPT condition. The results in Table 1 illustrate that the ChatTutor group ($M = 3.64$) scored significantly higher than the ChatGPT group ($M = 3.29$) on trust $t = 2.955$; $p = .002$. The ChatTutor group ($M = 3.65$) also scored significantly higher than the ChatGPT group ($M = 2.99$) on enjoyment $t = 3.304$; $p < .001$. Furthermore, the ChatTutor group ($M = 3.28$) scored significantly higher than the ChatGPT group ($M = 2.73$) on behavioral intentions $t = 2.188$; $p = .016$. However, the ChatTutor group ($M = 4.09$) scored only marginally higher than the ChatGPT group ($M = 3.82$) on self-efficacy $t = 1.661$; $p = .050$. Finally, there were no significant differences on self-efficacy between the ChatTutor and the teaching-as-usual conditions $t = .167$; $p = .434$.

Discussion

The first goal of this manuscript was to propose how GenAI could be used to support generative sense-making, providing a theoretical account of a learner-centered approach to using GenAI in education. Adopting an augmentation perspective, we use generative sense making research and theory to describe how GenAI can be used enhance lasting learning outcomes. More specifically, we propose how GenAI can be used to prime generative processing through enhancing personalized feedback when using teaching as a generative learning strategy. The second goal of this manuscript was to empirically investigate whether a theoretically informed GenAI chatbot (in this case, ChatTutor) could facilitate generative learning in education. Specifically, we aimed to test its impact on conceptual knowledge, self-efficacy, and trust immediately after a lesson, as well as its influence on conceptual knowledge, enjoyment, and behavioral intentions in a follow-up test four weeks later.

Empirical Contributions

The results of our study indicated that using ChatTutor resulted in significantly higher trust, enjoyment, and intentions to use the technology, and marginally higher self-efficacy compared to using ChatGPT. Here it is important to note that the students were unaware of the different conditions, as they were blind to the study's design and all used the ChatTutor interface. The difference was that the ChatTutor group obtained a theoretically informed GenAI learning intervention that builds on generative learning research and theory. Regarding conceptual knowledge the results indicate that the ChatTutor condition outperformed the ChatGPT condition in the immediate post-test but not in the delayed follow-up, and conversely the ChatTutor condition outperformed the teaching-as-usual condition in the follow-up but not the immediate post-test. The results related to conceptual knowledge are therefore mixed.

Delving into the mixed results related to conceptual knowledge in more depth, it is important to note that the groups only had 15 minutes on the follow-up generative activity task. This should be seen in the context of the entire course where learners were exposed to the topic of the Sternberg experiment through a seminar, lab activities and a specific lecture about the topic prior to the intervention. Therefore, it is likely that many students gained fundamental knowledge prior to engaging in the follow-up activity. While it would have been optimal to allow students more time to use ChatTutor we wanted to keep the amount of time across conditions constant as this has been highlighted as an important methodological consideration in previous studies (Brod, 2021). Furthermore, it is a strength that the study took place in an authentic learning setting where the value of GenAI for learning should ultimately be tested.

When delving into the literature on GLA there are some other factors to be aware of regarding our mixed results. Previous research on teaching as a GLA suggests that oral explanations may be more effective than written ones, as they can evoke a sense of social presence, leading to higher levels of arousal (Hoogerheide et al., 2019). A recent meta-analysis found a large positive effect of AI chatbots on learning outcomes (Wu & Yu, 2024). The authors recommend that future designers and educators further enhance these outcomes by incorporating human-like avatars, gamification elements, and emotional intelligence into AI chatbots. In this study we did not attempt to increase feelings of social presence through any of these methods which provide avenues for future research.

Furthermore, there is evidence that the GLA of teaching is most effective when the learning material is not available (Koh et al., 2018). This can be the case because it requires learners to retrieve information from long-term memory rather than relying on the lesson, aiding in knowledge consolidation and future accessibility (e.g., Roediger & Karpicke, 2006). In the current study the material was visually available for the students through the ChatTutor system. Therefore, it would have been possible to interact with ChatTutor by finding relevant passages in the text and paraphrasing them in the chat-box. Previous research has found that only students who generate an explanation exhibit improved understanding (Fiorella & Mayer, 2013; Hoogerheide et al., 2016). Therefore, future research should investigate the availability of learning content when using GenAI for GLS.

Theoretical Implications

The leading theoretical paradigms guiding research on learning have evolved over the past 100 years, reflecting paradigm shifts from response strengthening to information acquisition, knowledge construction, and knowledge co-construction (Mayer, 2012). Only time will tell if we are on the brink of a paradigm shift to a human-computer co-construction of knowledge. While it is unquestionable that existing educational psychology theory and research should guide and iteratively improve the development of GenAI-based learning interventions, as argued in this manuscript, a fundamental question remains: if and how educational psychology research must adapt to the new possibilities GenAI brings to education. The fundamental advantage of using GenAI lies in its ability to scale individualized learning experiences through realistic, dynamic conversations tailored to learners' needs. This approach shifts the focus from static, one-size-fits-all interventions to methods for delivering just-in-time support, aligning with advancements in learning analytics and self-regulated learning research (Azevedo & Gašević, 2019; Molenaar & Järvelä, 2014). In the context of generative sense-making, this could involve optimizing real-time feedback to support the creation and refinement of generative learning products, thereby enhancing engagement and understanding.

Practical Implications

One of the main challenges associated with the recent rapid advancements in GenAI is its disruptive impact. As formulated in Holmes et al. (2019), p 3: “If you can search, or have an intelligent agent find, anything, why learn anything? What is truly worth learning?” Hence, it is useful to discuss how GenAI-enhanced learning fits (or does not fit) with traditional views of learning outcomes and processes. The goal of most education is for learners to be able to apply what they have learned in a relevant context. This is referred to as meaningful learning or deep knowledge acquisition which is indicated by performance on transfer tests, which involves being able to use the learned material in new situations (Mayer, 2024). Kasneci et. al. (2023) highlight some of the opportunities and challenges of using GenAI in education. The opportunities include personalized learning experiences, and challenges include the need for critical thinking and strategies for fact checking because GenAI agents have a tendency to hallucinate and provide inaccurate information. In a society where knowledge is abundant and easily accessible through internet searches or conversations with GenAI chatbots, the goal of education cannot be to simply impart knowledge but rather to engage students to ask the right questions and develop the critical thinking skills needed to evaluate the information they receive. This would suggest the need to shift from remembering and focusing on higher order cognitive process dimensions such as understanding, applying, analyzing, evaluating, and creating from Anderson and Krathwohl's (2001) taxonomy. One challenge is that GenAI may make students regulate less (Yan et al., 2024). For example, while GenAI enhances the efficiency of information processing and retrieval, it poses a risk of fluency bias, where learners may overestimate their understanding due to the ease of processing information. Additionally, relying on GenAI for creative and problem-solving tasks could undermine these essential skills, fostering dependency and potentially hindering innovation and original thought (Rafner et al., 2023; Scneiderman, 2020). From a practical perspective this highlights the importance of thinking of GenAI as support rather than using it as a replacement as we have attempted to do in this manuscript.

Another important practical implication is the need to educate stakeholders on the differences between using generic GenAI tools like ChatGPT, educational GenAI, and educational GenAI designed with educational psychology theory. Our study results clearly show that an educational GenAI chatbot that builds on theory, such as ChatTutor, fosters significantly greater trust, enjoyment, and intentions to use the technology in the future compared to a generic GenAI system like ChatGPT. Another important practical consideration is the challenge teachers face in defining their role in the use of GenAI for learning. Yan et al. (2024) suggest incorporating human-in-the-loop elements, such as fostering partnerships among researchers, practitioners, and policymakers. This approach was used in our study, where we collaborated with instructors, researchers, and a GenAI development startup to design the ChatTutor intervention. Furthermore, the intervention included human-in-the-loop elements, as the ChatTutor system allowed learners to escalate a question to a human instructor if they encountered confusing or inaccurate information. This feature is crucial given that GenAI chatbots are known to 'hallucinate,' that is, provide false information.

Limitations and Future Research

Our study had several limitations. The primary limitation was the short duration learners had to interact with the GenAI chatbot, compared to the extensive time spent on the topic through a seminar, exercise course lectures, and lab activities. This setup was not ideal for measuring the tool's impact on learning outcomes but was necessary given the context of an authentic cognitive psychology course. Such an approach is typical of a value-added study, where the objective is to assess the additional benefit of an educational intervention. Future research should explore the value of using GenAI to facilitate GLS when students have more time to engage with a theoretically informed GenAI chatbot. Additionally, future studies should examine different age groups, as previous research has shown varying levels of GLS effectiveness across age groups (Broad, 2021).

Another limitation of our study concerns the reliability coefficients of some outcome measures, including conceptual knowledge and trust, which fell below acceptable levels. We prioritized content validity in our conceptual knowledge measure by including a broad range of items assessing different components of the learning material. However, future research would benefit from incorporating measures that are both valid and reliable. The goal of the learning intervention in this study was conceptual knowledge understanding. Future research could explore the value of theoretically informed GenAI for higher-order learning outcomes, such as applying, analyzing, evaluating, and creating, where relevant to learning interventions. Furthermore, this study focused on developing a GenAI intervention based on generative sense-making research and theory. Future research should explore how other educational psychology theories can inform the development of GenAI interventions, given the scarcity of such studies.

Conclusion

Generative Artificial Intelligence (GenAI) has captured significant public and scholarly attention and could challenge education by transforming traditional methods of teaching, learning, assessment, and administration. However, a recurring challenge with novel technologies lies in the tendency to implement and evaluate them from a technocentric perspective, often neglecting insights from educational psychology and learning science. This manuscript had two primary goals: first, to extend generative learning theory by proposing how GenAI can be utilized to support generative sense-making in education; and second, to empirically examine its impact on the outcomes of conceptual knowledge, self-efficacy, trust, enjoyment, and behavioral intentions. To achieve these objectives, we developed a theoretically informed GenAI chatbot (ChatTutor) designed to scaffold learning by prompting students to engage in the generative learning strategy of teaching a GenAI chatbot. Through a pre-registered experiment conducted in an authentic university setting, we compared ChatTutor to ChatGPT and a teaching-as-usual condition. The findings demonstrated that ChatTutor significantly enhanced trust, enjoyment, and behavioral intentions compared to ChatGPT, but not self-efficacy. Regarding learning outcomes, ChatTutor led to higher immediate conceptual knowledge scores than ChatGPT but did not outperform teaching-as-usual. In the delayed follow-up test, ChatTutor led to significantly higher conceptual knowledge than teaching-as-usual, but the differences with ChatGPT were no longer significant. This study underscores the importance of adapting educational psychology theories to harness the unique capabilities of GenAI in supporting meaningful learning. The findings contribute to the growing call for theory-driven, learner-centered integration of GenAI into education, moving beyond the initial fascination with the technology.

Declarations

Acknowledgements

We would like to thank Gustav B. Petersen, Benjamin B. Christensen, and Sidsel Drejer for helping develop the content of the lesson used in the experiment. Furthermore, we would like to thank Gustav B. Petersen, Marius Corson, Mikkel Marfelt and Kurt G. Nielsen for helping organize the experiment. Finally, we would like to thank Sidsel Drejer, Benjamin B. Christensen, Rubina F. Gogolu, Rebecca G. Lachmann, Alexander T. Ysbæk-Nielsen, Monique S. Damberg and Daniel Jensen who were the instructors who make the experiment possible.

References

1. Anderson, L. W., & Krathwohl, D. R. (2001). A Taxonomy for learning, teaching, and assesing. *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy*. Longman Publishing, 1–336.
2. Anthopic. (2024). *Claude's character*. Anthopic. https://doi.org/https://www.anthropic.com/research/claude-character
3. Azevedo, R., & Gašević, D. (2019). Analyzing multimodal multichannel data about self-regulated learning with advanced learning technologies: Issues and challenges. *Computers in Human Behavior*, *96*, 207–210. https://doi.org/10.1016/j.chb.2019.03.025
4. Bai, Y., Jones, A., Ndousse, K., Askell, A., Chen, A., DasSarma, N., Drain, D., Fort, S., Ganguli, D., Henighan, T., Joseph, N., Kadavath, S., Kernion, J., Conerly, T., El-Showk, S., Elhage, N., Hatfield-Dodds, Z., Hernandez, D., Hume, T., ... Kaplan, J. (2022). *Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback*, 1–74. https://doi.org/10.48550/arXiv.2204.05862
5. Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–214. https://doi.org/10.1037/0033-295X.84.2.191
6. Bandura, A. (2001). Social cognitive theory: An agentic perspective. In *Annual Review of Psychology, 52,* 1–26. https://doi.org/10.1146/annurev.psych.52.1.1
7. Benbasat, I., & Wang, W. (2005). Trust In and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems*, *6*(3), 72–101. https://doi.org/10.17705/1jais.00065
8. Bereiter, C., & Scardamalia, M. (2013). The psychology of written composition. In *The Psychology of Written Composition*, 1–389. https://doi.org/10.4324/9780203812310
9. Berlyne, D. E. (2006). Conflict, arousal, and curiosity. In *Conflict, arousal, and curiosity,* 1–366. https://doi.org/10.1037/11164-000
10. Berthold, K., & Renkl, A. (2009). Instructional Aids to Support a Conceptual Understanding of Multiple Representations. *Journal of Educational Psychology*, *101*(1), 70–87. https://doi.org/10.1037/a0013247
11. Berthold, K., Eysink, T. H., & Renkl, A. (2009). Assisting selfexplanation prompts are more effective than open prompts when learning fromm multiple representations. *Instructional Science, 37,* 345–363. https://doi.org/10.1007/s11251-008-9051-z
12. Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing Self-Explanation: a Meta-Analysis. *Educational Psychology Review*, *30*(3), 703–725. https://doi.org/10.1007/s10648-018-9434-x
13. Brennan, K. (2015). Beyond technocentrism: Supporting constructionism in the classroom. *Constructivist Foundations*, *10*(3), 289–286. https://constructivist.info/10/3/289.brennan.pdf
14. Brod, G. (2021). Generative Learning: Which Strategies for What Age? In *Educational Psychology Review*. *33*(4), 1295–1318. https://doi.org/10.1007/s10648-020-09571-9
15. Burkhart, C., Lachner, A., & Nückles, M. (2021). Using Spatial Contiguity and Signaling to Optimize Visual Feedback on Students' Written Explanations. *Journal of Educational Psychology*, *113*(5), 998–1023. https://doi.org/10.1037/edu0000607
16. Castro-Alonso, J. C., de Koning, B. B., Fiorella, L., & Paas, F. (2021). Five Strategies for Optimizing Instructional Materials: Instructor- and Learner-Managed Cognitive Load. In *Educational Psychology Review*, *33*(4), 1379–1407. https://doi.org/10.1007/s10648-021-09606-9
17. Chandler, P. (2009). Dynamic visualisations and hypermedia: Beyond the "Wow" factor. *Computers in Human Behavior*, *25*(2), 389–392. https://doi.org/10.1016/j.chb.2008.12.018
18. Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. https://doi.org/10.1016/0364-0213(89)90002-5
19. Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments, 32*(7), 1–17. https://doi.org/10.1080/10494820.2023.2172044
20. Chow, A. (2023). How ChatGPT Managed to Grow Faster Than TikTok or Instagram. *TIME*. Retrieved https://time.com/6253615/chatgpt-fastest-growing/
21. Cuéllar, M. F., Larsen, B., Lee, Y. S., & Webb, M. (2024). Does Information About AI Regulation Change Manager Evaluation of Ethical Concerns and Intent to Adopt AI? *The Journal of Law, Economics, and Organization*, *40*(1), 34–75. https://doi.org/10.1093/jleo/ewac004
22. Dai, J., Zhang, X., & Wang, C. (2024). A meta-analysis of learners' continuance intention toward online education platforms. *Education and Information Technologies, 29,* 1–36. https://doi.org/10.1007/s10639-024-12654-7
23. Dargue, N., Sweller, N., & Jones, M. P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765–784. https://doi.org/10.1037/bul0000202
24. Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers and Education*, *210,* 1–18. https://doi.org/10.1016/j.compedu.2023.104967

25. Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Applied Social Psychology*, *22*(14), 1111–1132. https://doi.org/10.1111/j.1559-1816.1992.tb00945.x
26. De Jong, T. (2006). Technological advances in inquiry learning. In *Science, 312*(5773), 532–533. https://doi.org/10.1126/science.1127750
27. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *proceedings of NAACL HLT 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies,* (pp. 1–16)*. https://doi.org/10.48550/arXiv.1810.04805*
28. Dillenbourg, P. (1999). Collaborative Learning: Cognitive and Computational Approaches. Advances in Learning and Instruction Series. *Elsevier Science*, 1–246. https://eric.ed.gov/?id=ED437928
29. Edwards, B. (2023). OpenAI's GPT-4 exhibits "human-level performance" on professional benchmarks. *Ars Technica.* Retrieved from: https://arstechnica.com/information-technology/2023/03/openai-announces-gpt-4-its-next-generation-ai-language-model/
30. Eshuis, E. H., ter Vrugte, J., Anjewierden, A., & de Jong, T. (2022). Expert examples and prompted reflection in learning with self-generated concept maps. *Journal of Computer Assisted Learning*, *38*(2), 350–365. https://doi.org/10.1111/jcal.12615
31. Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/BF03193146
32. Fenn, J., & Blosch, M. (2020). *Understanding Gartner's Hype Cycles.* Gartner Research. https://www.gartner.com/en/documents/3887767
33. Fiorella, L. (2023). Making Sense of Generative Learning. In *Educational Psychology Review, 35*(2), 1–42. https://doi.org/10.1007/s10648-023-09769-7
34. Fiorella, L., & Kuhlmann, S. (2020). Creating Drawings Enhances Learning by Teaching. *Journal of Educational Psychology*, *112*(4), 811–822. https://doi.org/10.1037/edu0000392
35. Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology, 38*(4), 281–288. https://doi.org/10.1016/j.cedpsych.2013.06.001
36. Fiorella, L., & Mayer, R. E. (2015). Learning as a generative activity: Eight learning strategies that promote understanding. 1–218. https://doi.org/10.1017/CBO9781107707085
37. Fiorella, L., & Mayer, R. E. (2016). Eight Ways to Promote Generative Learning. In *Educational Psychology Review*, *28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9
38. Fiorella, L., & Mayer, R. E. (2017). Spontaneous spatial strategy use in learning from scientific text. *Contemporary Educational Psychology*, *49*, 66–79. https://doi.org/10.1016/j.cedpsych.2017.01.002
39. Fiorella, L., & Mayer, R. E. (2022). The Generative Activity Principle in Multimedia Learning, 339–350. In *The Cambridge Handbook of Multimedia Learning*. https://doi.org/10.1017/9781108894333.036
40. Fiorella, L., Yoon, S. Y., Atit, K., Power, J. R., Panther, G., Sorby, S., Uttal, D. H., & Veurink, N. (2021). Validation of the Mathematics Motivation Questionnaire (MMQ) for secondary school students. *International Journal of STEM Education*, *8*(1), 1–14. https://doi.org/10.1186/s40594-021-00307-x
41. Fiorella, L., & Zhang, Q. (2018). Drawing Boundary Conditions for Learning by Drawing. In *Educational Psychology Review, 30*(3), 1115–1137. https://doi.org/10.1007/s10648-018-9444-8
42. Fischer, G., Lundin, J., & Lindberg, J. O. (2020). Rethinking and reinventing learning, education and collaboration in the digital age—from creating technologies to transforming cultures. *International Journal of Information and Learning Technology*, *37*(5), 241–252. https://doi.org/10.1108/IJILT-04-2020-0051
43. Goldin-Meadow, S., Cook, S. W., & Mitchell, Z. A. (2009). Gesturing gives children new ideas about math. *Psychological Science*, *20*(3), 267–272. https://doi.org/10.1111/j.1467-9280.2009.02297.x
44. Gulati, S., Sousa, S., & Lamas, D. (2019). Design, development and evaluation of a human-computer trust scale. *Behaviour and Information Technology*, *38*(10), 1004–1015. https://doi.org/10.1080/0144929X.2019.1656779
45. Helmore, E. (2023). We are a little bit scared': OpenAI CEO warns of risks of artificial intelligence. *TheGuardian*. Retrieved from: https://www.theguardian.com/technology/2023/mar/17/openai-sam-altman-artificial-intelligence-warning-gpt4
46. Hennessy, S., Cukurova, M., Lewin, C., Mavrikis, M., & Major, L. (2024). BJET Editorial 2024: A call for research rigour. In *British Journal of Educational Technology*, *55*(1), 5–9. https://doi.org/10.1111/bjet.13426
47. Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). In *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368
48. Hoogerheide, V., Deijkers, L., Loyens, S. M., Heijltjes, A., & van Gog, T. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. *Contemporary Educational Psychology, 44,* 95–106. https://doi.org/10.1016/j.cedpsych.2016.02.005
49. Hoogerheide, V., Renkl, A., Fiorella, L., Paas, F., & Van Gog, T. (2019). Enhancing example-based learning: Teaching on video increases arousal and improves problem-solving performance. *Journal of Educational Psychology, 111*(1), 45–56. https://doi.org/10.1037/edu0000272

50. Holmes, W., Fadel, C., & Bialik, M. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. *Journal of Computer Assisted Learning*, *14*(4). https://www.researchgate.net/publication/332180327_Artificial_Intelligence_in_Education_Promise_and_Implications_for_Teaching_
51. Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533
52. Huang, F., & Zou, B. (2024). English speaking with artificial intelligence (AI): The roles of enjoyment, willingness to communicate with AI, and innovativeness. *Computers in Human Behavior*, *159,* 2–8. https://doi.org/10.1016/j.chb.2024.108355
53. IBM Corp. (2023). IBM SPSS Statistics for Windows. https://doi.org/https://www.ibm.com/products/spss-statistics
54. Järvelä, S., Nguyen, A., Vuorenmaa, E., Malmberg, J., & Järvenoja, H. (2023). Predicting regulatory activities for socially shared regulation to optimize collaborative learning. *Computers in Human Behavior*, *144,* 1–10. https://doi.org/10.1016/j.chb.2023.107737
55. Johnson, C. I., & Marrafno, M. D. (2022). The feedback principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 286– 295). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.023
56. 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. In *Learning and Individual Differences*, *103,* 1–9. https://doi.org/10.1016/j.lindif.2023.102274
57. Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, *3,* 1–22. https://doi.org/10.1016/j.caeai.2022.100074
58. King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology, 90*, 134–152. https://doi.org/10.1037/0022-0663.90.1.134
59. Kobayashi, K. (2019). Learning by preparing-to-teach and teaching: A meta-analysis. *Japanese Psychological Research, 61*(3), 192–203. https://doi.org/10.1111/jpr.12221
60. Koh, A. W. L., Lee, S. C., & Lim, S. W. H. (2018). The learning benefits of teaching: A retrieval practice hypothesis. *Applied Cognitive Psychology, 32*(3), 401–410. https://doi.org/10.1002/acp.3410
61. Lachner, A., Jacob, L., & Hoogerheide, V. (2021). Learning by writing explanations: Is explaining to a fictitious student more effective than self-explaining? *Learning and Instruction*, *74,* 1–13. https://doi.org/10.1016/j.learninstruc.2020.101438
62. Lachner, A., & Neuburg, C. (2019). Learning by writing explanations: computer-based feedback about the explanatory cohesion enhances students' transfer. *Instructional Science*, *47*(1), 19–37. https://doi.org/10.1007/s11251-018-9470-4
63. Lee, S. Y., & Choi, J. (2017). Enhancing user experience with conversational agent for movie recommendation: Effects of self-disclosure and reciprocity. *International Journal of Human Computer Studies*, *103,* 95–105. https://doi.org/10.1016/j.ijhcs.2017.02.005
64. Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*(10), 464-470. https://doi.org/10.1016/j.tics.2006.08.004
65. Lorenz, P., Perset, K., & Berryhill, J. (2023). Initial policy considerations for generative artificial intelligence. *OECD Artificial intelligence papers*, *1,* 1–40. Retrieved from: https://www.oecd.org/en/publications/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en.html
66. Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review, 33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2
67. Manalo, E., Uesaka, Y., & Chinn, C. A. (2017). Promoting Spontaneous Use of Learning and Reasoning Strategies. In *Promoting Spontaneous Use of Learning and Reasoning Strategies,* 1–166. https://doi.org/10.4324/9781315564029
68. Mayer, R. E. (1996). Learning strategies for making sense out of expository text: the soi model for guiding three cognitive processes in knowledge construction. *Educational Psychology Review*, *8*(4), 357–371. https://doi.org/10.1007/BF01463939
69. Mayer, R. E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist*, *59*(1), 14–19 . https://doi.org/10.1037/0003-066x.59.1.14
70. Mayer, R. E. (2012). Getting Started on the Road to Applying the Science of Learning. In *Applied Cognitive Psychology*, *26*(2), 330–331. https://doi.org/10.1002/acp.1829
71. Mayer, R. E. (2014). *Cognitive Theory of Multimedia Learning* (2nd ed). Cambridge University Press, 43–71. https://doi.org/10.1017/CBO9781139547369.005
72. Mayer, R. E. (2024). The Past, Present, and Future of the Cognitive Theory of Multimedia Learning. *Educational Psychology Review*, *36*(1), 1–25. https://doi.org/10.1007/s10648-023-09842-1
73. McNamara, D. S. (2017). Self-Explanation and Reading Strategy Training (SERT) Improves Low-Knowledge Students' Science Course Performance. *Discourse Processes*, *54*(7), 479–492. https://doi.org/10.1080/0163853X.2015.1101328

74. Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, *57*(4), 632–645. https://doi.org/10.1111/ejed.12527
75. Molenaar, I., & Järvelä, S. (2014). Sequential and temporal characteristics of self and socially regulated learning. *Metacognition and Learning*, *9*, 75–85. https://doi.org/10.1007/s11409-014-9114-2
76. Murayama, K. (2022). A reward-learning framework of knowledge acquisition: An integrated account of curiosity, interest, and intrinsic–extrinsic rewards. *Psychological Review*, *129*(1), 175. https://doi.org/10.1037/rev0000349
77. Nie, A., Chandak, Y., Suzara, M., Ali, M., Woodrow, J., Peng, M., Sahami, M., Brunskill, E., & Piech, C. (2024). The GPT Surprise: Offering Large Language Model Chat in a Massive Coding Class Reduced Engagement but Increased Adopters Exam Performances, *arXiv,* 1–33. http://arxiv.org/abs/2407.09975
78. Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, *14,* 47–61. https://doi.org/10.1016/j.edurev.2015.02.003
79. Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review*, *18*, 315–341. https://doi.org/10.1007/s10648-006-9029-9
80. Pintrich, P. R. R., Smith, D., Garcia, T., & McKeachie, W. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). *ERIC,* 1–76. Retrieved from: https://files.eric.ed.gov/fulltext/ED338122.pdf
81. Ponce, H. R., Mayer, R. E., Loyola, M. S., & López, M. J. (2020). Study Activities That Foster Generative Learning: Notetaking, Graphic Organizer, and Questioning. *Journal of Educational Computing Research*, *58*(2), 1–22. https://doi.org/10.1177/0735633119865554
82. Qadir, J. (2023). Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education. In *IEEE Global Engineering Education Conference, EDUCON*, 1–9. https://doi.org/10.1109/EDUCON54358.2023.10125121
83. Rafner, J., Beaty, R. E., Kaufman, J. C., Lubart, T., & Sherson, J. (2023). Creativity in the age of generative AI. *Nature Human Behaviour*, *7*(11), 1836–1838. https://doi.org/10.1038/s41562-023-01751-1
84. Rellensmann, J., Schukajlow, S., Blomberg, J., & Leopold, C. (2022). Effects of drawing instructions and strategic knowledge on mathematical modeling performance: Mediated by the use of the drawing strategy. *Applied Cognitive Psychology*, *36*(2), 402–417. https://doi.org/10.1002/acp.3930
85. Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, *21*(1), 1–29. https://doi.org/10.1207/s15516709cog2101_1
86. Renkl, A., & Scheiter, K. (2017). Studying Visual Displays: How to Instructionally Support Learning. *Educational Psychology Review*, *29*(3), 599–621. https://doi.org/10.1007/s10648-015-9340-4
87. Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x
88. Roelle, J., & Berthold, K. (2017). Effects of incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction*, *49,* 142–156. https://doi.org/10.1016/j.learninstruc.2017.01.008
89. Roschelle, J., Lester, J., Fusco, J., Safir, A., Johnstun, K., Trettin, S., Chhin, C., Metz, E., Digital Promise colleagues, O., Cator, K., Means, B., Bellin, M., & Van Ostrand, K. (2020). AI and the Future of Learning: Expert Panel Report. *Digital Promise,* 1–27. Retrived from: https://circls.org/wp-content/uploads/2020/11/CIRCLS-AI-Report-Nov2020.pdf
90. Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, *77*(4), 534–574. https://doi.org/10.3102/0034654307309920
91. Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, *39*(4), 275–286. https://doi.org/10.1016/j.cedpsych.2014.07.003
92. Shneiderman, B. (2020). Human-centered artificial intelligence: reliable, safe & trustworthy. *Int. J. Hum. Comput. Interact. 36,* 495–504. https://doi.org/10.1080/10447318.2020.1741118
93. Schramowski, P., Turan, C., Andersen, N., Rothkopf, C. A., & Kersting, K. (2022). Large pre-trained language models contain human-like biases of what is right and wrong to do. *Nature Machine Intelligence*, *4*(3), 258–268. https://doi.org/10.1038/s42256-022-00458-8
94. Schukajlow, S., Krawitz, J., Kanefke, J., & Rakoczy, K. (2022). Interest and performance in solving open modeling problems and closed real-world problems. *PME, 3.* Retrieved from: https://ivv5hpp.uni-muenster.de/u/sschu_12/pdf/Publikationen/Schukajlow_etal_2022_PME45.pdf
95. Schwamborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a Generative Activity and Drawing as a Prognostic Activity. *Journal of Educational Psychology*, *102*(4), 872–879. https://doi.org/10.1037/a0019640
96. Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist*, *57*(1), 31–54. https://doi.org/10.1080/00461520.2021.1939700
97. Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*(3736), 652–654. https://doi.org/10.1126/science.153.3736.652

98. Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. American Scientist, 57(4), 421–457. Retrieved from: https://www.sas.upenn.edu/~saul/am.scientist69.pdf
99. Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. In *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5
00. Taecharungroj, V. (2023). "What Can ChatGPT Do?" Analyzing Early Reactions to the Innovative AI Chatbot on Twitter. *Big Data and Cognitive Computing*, *7*(1), 1–10. https://doi.org/10.3390/bdcc7010035
01. Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). LLAMA: Open and Efficient Foundation Language Models. *ArXi,* 1–17. https://doi.org/https://doi.org/10.48550/arXiv.2302.13971
02. van de Pol, J., van Loon, M., van Gog, T., Braumann, S., & de Bruin, A. (2020). Mapping and Drawing to Improve Students' and Teachers' Monitoring and Regulation of Students' Learning from Text: Current Findings and Future Directions. In *Educational Psychology Review*, *32*(4), 951–977. https://doi.org/10.1007/s10648-020-09560-y
03. Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*(1), 129–977. https://doi.org/10.1037/0022-0663.93.1.129
04. Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, *39*(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x
05. Vygotsky, L. S. (1978). Mind and Society: The Development of Higher Psychological Processes. In *Harvard University Press,* 1–174. https://doi.org/10.2307/j.ctvjf9vz4
06. Willingham, D. T. (2008). Critical Thinking: Why Is It So Hard to Teach? *Arts Education Policy Review*, *109*(4), 21–32. https://doi.org/10.3200/AEPR.109.4.21-32
07. Wittrock, M. C. (1989). Generative Processes of Comprehension. *Educational Psychologist*, *24*(4), 345–275. https://doi.org/10.1207/s15326985ep2404_2
08. Wittrock, M. C., & Farley, F. (2010). Learning as a generative process. *Educational Psychologist*, *45*(1), 40–45. https://doi.org/10.1080/00461520903433554
09. Wittwer, J., & Renkl, A. (2010). How Effective are Instructional Explanations in Example-Based Learning? A Meta-Analytic Review. In *Educational Psychology Review, 22*(4), 393–409. https://doi.org/10.1007/s10648-010-9136-5
10. Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33.
11. https://doi.org/10.1111/bjet.13334
12. Yan, L., Greiff, S., Teuber, Z., & Gašević, D. (2024). Promises and challenges of generative artificial intelligence for human learning. *Nature Human Behavior, 8,* 1839–1850. https://doi.org/10.1038/s41562-024-02004-5
13. Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic scoping review. In *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370
14. Zhang, Q., & Fiorella, L. (2023). An integrated model of learning from errors. *Educational Psychologist*, *58*(1), 18–34. https://doi.org/10.1080/00461520.2022.2149525

Table

Table 1 is available in the Supplementary Files section

Figures

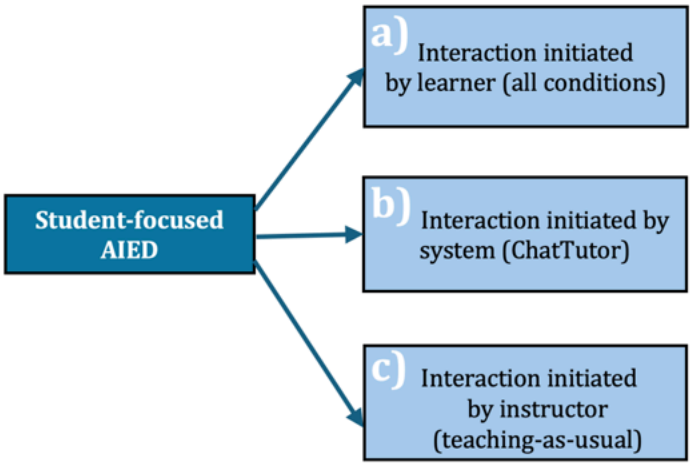

Figure 1

Framework for the development of a student-focused GenAI educational application where interaction is initiated by a) the learner, b) the AIEd system, and c) the instructor, adapted from the taxonomy developed by Holmes and Tuomi (2022).

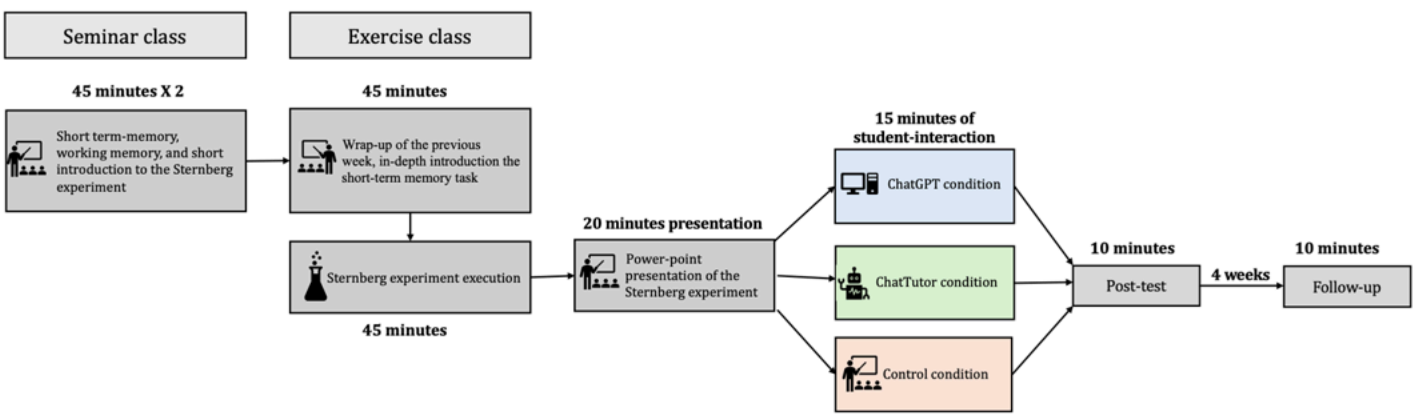

Figure 2

Figure of experimental procedure

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplamentaryMaterial.docx
- Table1.docx