



Leveraging the Potential of Large Language Models in Education Through Playful and Game-Based Learning

Stefan E. Huber¹ · Kristian Kiili² · Steve Nebel³ · Richard M. Ryan^{4,5} · Michael Sailer⁶ · Manuel Ninaus^{1,7}

Accepted: 9 February 2024
© The Author(s) 2024

Abstract

This perspective piece explores the transformative potential and associated challenges of large language models (LLMs) in education and how those challenges might be addressed utilizing playful and game-based learning. While providing many opportunities, the stochastic elements incorporated in how present LLMs process text, requires domain expertise for a critical evaluation and responsible use of the generated output. Yet, due to their low opportunity cost, LLMs in education may pose some risk of over-reliance, potentially and unintendedly limiting the development of such expertise. Education is thus faced with the challenge of preserving reliable expertise development while not losing out on emergent opportunities. To address this challenge, we first propose a playful approach focusing on skill practice and human judgment. Drawing from game-based learning research, we then go beyond this playful account by reflecting on the potential of well-designed games to foster a willingness to practice, and thus nurturing domain-specific expertise. We finally give some perspective on how a new pedagogy of learning with AI might utilize LLMs for learning by generating games and gamifying learning materials, leveraging the full potential of human-AI interaction in education.

Keywords Large language models · Generative artificial intelligence · Education · Playful learning · Gamification · Game-based learning

Introduction

Large language models (LLMs) and their recently increased accessibility via chatbots like ChatGPT (OpenAI, 2023), Bard (Google, 2023), or Bing Chat (Microsoft, 2023) provide both new opportunities and challenges for education. On the one hand, they legitimately promise effective ways to assist with many tasks involved in both teaching and learning (Bernabei et al., 2023; Kohnke et al., 2023), to provide scalable, personalized learning material (Abd-alrazaq et al., 2023; Sallam, 2023),

Extended author information available on the last page of the article

and thus easy and scalable opportunities for exercise (Kasneci et al., 2023). On the other hand, they come with the educational challenge to avoid becoming overly or naïvely reliant on their support (Abd-alrazaq et al., 2023; Bernabei et al., 2023; Kasneci et al., 2023; Kohnke et al., 2023; Shue et al., 2023; Zhu et al., 2023), and thus to prevent inadvertently adopting inherent biases (Abd-alrazaq et al., 2023; Bernabei et al., 2023; Dwivedi et al., 2023; Kasneci et al., 2023; Zhu et al., 2023) or losing out on opportunities for reflection and practice for developing domain expertise and judgment competence (Dwivedi et al., 2023; Krügel et al., 2023). These are, however, especially needed for the responsible use of present LLMs, because, due to their inherent random mechanisms utilized during text generation (Wolfram, 2023), mistakes or fabricated information cannot be entirely ruled out. Hence, at least for the time being, the output generated by LLMs definitely requires domain expertise for critical revision and evaluation. Education thus finds itself currently faced with the challenge to find an appropriate balance between seizing new and welcome opportunities and protecting against inadvertent risks of losing out on the development of required expertise at the same time.

In this perspective piece, we propose that—in a first step—a more exploratory, playful approach towards the use of LLMs may help with finding such an appropriate balance. Such an approach has already been utilized in the form of prompt engineering in various domains (Oppenlaender et al., 2023; Polak & Morgan, 2023; Short & Short, 2023; Shue et al., 2023; Wang et al., 2023; White et al., 2023; Zhu et al., 2023). Beyond those accounts, we further suggest that—in a second step—going the full way to a game-based education could eventually provide a new pedagogy of learning with artificial intelligence (AI) leveraging the full potential within a well-balanced cooperation between human and machine intelligence. We further argue that this second step allows utilization of LLMs for devising appropriate game-based learning environments, such that LLMs may eventually serve to overcome exactly those challenges they pose for education in the first place.

To serve a systematic development of our arguments, the article is organized as follows: first, we briefly illustrate both opportunities and challenges posed by the usage of LLMs in educational contexts. In a second section, we argue how a more playful approach to the usage of LLMs in education may already help to resolve some of the tension between opportunities, challenges, and risks. In a final section, we outline our proposition how game-based learning can extend the limits of said playful approaches, paving the way for a prolific co-operation between human and artificial intelligence in education.

LLMs in Education—Opportunities and Challenges

Generally, LLMs are a recently developed form of AI (i.e., algorithms historically devised to mimic, extend, or replace parts of human cognition or behavior). More specifically, they are a form of *generative* AI, representing algorithms capable of generating new media like images or text.

Recent LLMs (like those provided via ChatGPT) use large datasets of text in conjunction with artificial neural networks with billions of parameters to process and

generate text. Chat-like interfaces allow the user to obtain human-like responses in conversational style upon entering arbitrary prompts. While earlier language models like Wordtune, Paperpal, or Generate (Hutson, 2022) could help writers restructure a sentence, more recent versions like ChatGPT can help with devising entire manuscripts, providing feedback, finding limitations (Zimmerman, 2023), or devising specialized text like computer code (Shue et al., 2023).

The essential core principles have, however, remained similar (Wolfram, 2023): the computation of likely continuations of the user-provided prompt based on identified relations between text elements in the vast amount of training data. An important ingredient in the computation is the fact that not always the most likely, but sometimes a less likely continuation is chosen. While this serves the impression of an especially spontaneous, human-like, fluently emergent text, it also is the reason why the information provided by present LLMs can be misleading or erroneous and thus requires continuous supervision and critical evaluation.

Opportunities

Given their capabilities, LLMs provide a wide range of opportunities for education (Kasneci et al., 2023). LLMs can assist with management tasks (e.g., development of teaching units, curricula, or personalized study plans), with assessment and evaluation, and with program monitoring and review (Abd-alrazaq et al., 2023). They can take the roles of content providers (Abd-alrazaq et al., 2023; Jeon & Lee, 2023; Sarsa et al., 2022), temporary interlocutors, teaching assistants, and evaluators (Jeon & Lee, 2023). They can assist with writing tasks of both teachers and learners (Bernabei et al., 2023), regarding not only content creation, but also basic information retrieval (Zhu et al., 2023) and literature review (Abd-alrazaq et al., 2023).

LLMs can further assist teachers in orchestrating a continuously growing plethora of teaching resources, making the teachers' resources (bound to developing and revising learning material in earlier times) more available for designing creative, well-organized, and engaging lessons (Jeon & Lee, 2023). They enable personalized learning (Abd-alrazaq et al., 2023; Sallam, 2023) and may benefit learners' understanding of topics (Bernabei et al., 2023; Sarsa et al., 2022; Zhu et al., 2023). If used carefully, they can enhance critical thinking and problem-based learning (Bernabei et al., 2023; Sallam, 2023; Shue et al., 2023), emphasize the role of students as active investigators, and raise ethical awareness regarding the use of AI (Jeon & Lee, 2023).

Challenges

However, careful use of LLMs also presents a challenge to both teachers and learners (Kasneci et al., 2023). This is related to a variety of shortcomings of LLMs that have not yet been entirely resolved. These include the possibility of mistakes or fabricated information; the lack of recent, state-of-the-art, domain knowledge; the lack of originality; inherent (social or gender) biases; various ethical and legal issues like

copyright, plagiarism, and false citations; lacks of transparency and accountability; cybersecurity issues; and the risk of infodemics (Sallam, 2023; Zhu et al., 2023).

In contrast to pocket calculators, present LLMs are not designed to yield reliably the same deterministic output upon the same given prompt. A stochastic element in generation of such output is a part of how and why they work so astonishingly well in producing seemingly human-like responses (Wolfram, 2023). This, however, has also the consequence that their output definitely requires critical evaluation and careful revision by domain experts (Ali et al., 2023; Biswas, 2023; Hosseini et al., 2023; Howard et al., 2023; Kasneci et al., 2023; Mogali, 2023; Salvagno et al., 2023; Van Dis et al., 2023; Zhu et al., 2023). Especially when it is about decisions that should guide human action, the support provided by LLMs should be supervised by human expertise (Molenaar, 2021).

Expertise as a Crucial Factor in Human-AI Systems

This resonates well with the general assertion that the quality of decisions by human-AI systems depends crucially on the human expertise within such systems (Ninaus & Sailer, 2022). However, both the development and preservation of expertise require practicing domain-specific problem-solving capabilities (Elvira et al., 2017; Tynjälä, 2008; Tynjälä et al., 2006).

As novices advance from easier to more difficult problems, they continuously engage in three learning processes. First, they transform conceptual knowledge into experiential knowledge when, for instance, applying general concepts to specific problems in particular contexts. Second, they also need to explicate experiential into conceptual knowledge to, for instance, make tacit knowledge (Patterson et al., 2010) accessible to other people as well as to metacognitive processes like reflection. Reflecting on experiential and conceptual knowledge finally allows for improving problem-solving strategies, further supports the transfer of tacit to explicable knowledge, and facilitates the development of learning strategies, metacognitive, and self-regulatory skills (Elvira et al., 2017).

All three processes have in common that continuous practice in integrating conceptual, experiential, and self-reflective knowledge during problem-solving utilizes already existing expertise and contributes to its further development. Although modern theories on expertise acknowledge that many factors besides practice contribute to expertise development (Hambrick et al., 2016), they do not deny the relevance or even necessity of (deliberate) practice (Campitelli & Gobet, 2011; Ericsson et al., 1993; Hambrick et al., 2014).

Interaction Between Use of LLMs and Expertise Development

In formal education, which lays the foundations for the development of expertise, practice sometimes requires that learners engage in effortful or even strenuous tasks. That is, learners need to regulate their attention and efforts toward a task that might be associated with aversive feelings and also to resist engaging in more pleasurable activities (Kurzban et al., 2013; Miller et al., 2012).

However, the convenience and low opportunity cost that LLMs bring for certain tasks, bears the risk of over-reliance (Kasneci et al., 2023) or over-trust (Morris et al., 2023), which has also been recognized as a hindrance for critical thinking (Shue et al., 2023), learning, and reflection (Zhu et al., 2023). In addition to that, learners (and sometimes also teachers) can feel tempted by the authoritative nature of the responses to take them at face value without critically evaluating and processing them further (Kohnke et al., 2023). Lastly, learners can be tempted to outsource the activity. While such outsourcing might be appropriate for tasks that are merely means to an end, it becomes problematic when tasks represent essential learning opportunities for skills that a person should have even without AI support (Salomon et al., 1991). Over-reliance on LLMs in educational contexts is thus associated with some risk of losing out on essential ingredients for the development and preservation of expertise, potentially and inadvertently providing also a risk of deskilling (Morris et al., 2023), and consequentially of automation bias, reduced human autonomy and judgment competence (Dwivedi et al., 2023; Deutscher Ethikrat, 2023).

It is important, however, to note that an eventual shift in what is considered an essential skill is not problematic per se. As with every new useful tool, LLMs also bring about a shift in what is considered essential expertise. While in earlier times, doing a statistical analysis might have involved manually integrating a normal curve to determine a p -value, this would hardly suggest that a social scientist not knowing anymore how to do this has not developed any statistical expertise (we thank the anonymous reviewer for providing this example). The advent of the digital computer has changed the outline of the skill set determining the meaning of statistical expertise.

Ongoing developments of generative AI technology like retrieval augmented generation, improving on both factual reliability and timeliness of responses provided by LLMs (Gao et al., 2023), are likely to push the boundaries of what kind of expertise may be called essential even further. The critical point remains that high-quality decisions of human-AI systems presuppose *some* human expertise (Ninaus & Sailer, 2022). And it is difficult to judge in advance which skill sets will remain essential in the future. As Dwivedi et al. (2023) note, we as educators must ask ourselves first: which skills are still needed? Once these are identified, a second question remains: how can we devise new, appropriate ways of developing and practicing these skills in a new pedagogy of learning with AI?

Banning LLMs?

One response addressing this challenge are calls for more closely regulating the use of LLMs, ranging from simply requiring disclosure (Stokel-Walker, 2023) over adaption of examination procedures (Dwivedi et al., 2023) to complete bans (Johnson, 2023; Rosenzweig-Ziff, 2023). Yet attempts at external control face at least one very pragmatic issue: It can be difficult, if possible at all, to distinguish between human- and AI-produced material (Ariyaratne et al., 2023; Dunn et al., 2023; Else, 2023). Although tools are developed that allow (at least temporarily) AI-support detection to some extent (Bernabei et al., 2023; Else, 2023), we also think that

research and higher education needs to devise ways to use LLMs ethically, transparently, and responsibly (Van Dis et al., 2023). Furthermore, “it makes no sense to ban the technology for students that will live in a world where this technology will play a major role” (Dwivedi et al., 2023, p. 9).

A completely different response to the outlined challenge originates long before the most recent advent of AI in the form of LLMs. It involves a more playful stance towards the new possibilities that come with new technology.

On Playful Approaches to Integrate New Technology in Education

As early as in the 1960s, Papert (1980) developed a pedagogical approach which allowed to utilize computers to facilitate children’s understanding of geometry. However, instead of thinking of ways to use computers just as providers of more sophisticated, digital teaching or learning material, children were enabled to build up their geometrical understanding by providing them with a tool to let computers do something meaningful to them. For this purpose, the programming language Logo was developed (Papert, 1980) which allowed children to control the movement of a virtual turtle which left behind lines as it moved over the screen. By understanding how to draw geometrical shapes by controlling the turtle, and further, how simple geometrical shapes constitute more complex images, a gradually improving understanding of geometry allowed the children to draw more beautiful and complex images. Playful experimentation with the Logo language allowed to build up experiential knowledge by applying basic, conceptual knowledge of how to draw squares, triangles, and so forth. At the same time, purposeful drawing of more complex, composite objects (like a house with a door, windows, and a roof) required translating experiential knowledge into conceptual knowledge by the necessity to provide specific commands. Learning by purposive doing and by engaging in discovery via the natural processes of trial and error would further provide ample opportunity to reflect on both, experiential and conceptual knowledge to further improve drawing capabilities and thus, understanding geometry. Papert’s pedagogical approach (1980), hence, naturally nurtured all three learning processes involved in developing expertise (Elvira et al., 2017; Tynjälä, 2008; Tynjälä et al., 2006). Not only became children able to produce images and experiences of meaning for themselves, but they did so just inasmuch as they improved in their geometrical understanding, programming capabilities, and computational thinking. Furthermore, new technology, i.e., the digital computer, which could have just been programmed to do the same geometrical operations much more efficiently, was instead utilized to promote education (Papert, 1980).

Yet, why did Papert come up with his playful, constructionist approach to learning in the first place? In fact, he was inspired by constructivist theory of how children construct new schemas by interacting with their environment (Piaget, 1962). In Piaget’s theory of cognitive development (1962), play facilitates children’s cognitive development by activating basic units for organizing knowledge and behavior, known as schemas. Play allows both the practice of existing schemas, and thus of

existing skills and knowledge, and the development of new ones by combining elements of existing ones in ways that transcend existing knowledge.

Social development theory (Vygotsky, 1967), scrutinizing also the developmental importance of play, adds the notion that the crucial point of play for learning is its capability to provide children with opportunities to explore outcomes beyond their current abilities. Play allows players to experience and simulate potential outcomes without the real-life costs (Homer et al., 2020). It allows to probe their capabilities, and by that, it allows them to grow beyond their current limitations. Although highlighting somewhat different aspects, both theories of play highlight their potential for facilitating learning and development.

More recently, research within *self-determination theory* (SDT; Ryan & Deci, 2017) has specifically highlighted the importance of intrinsic motivation, the enjoyment of the activity itself, as critical to learning across development (Reeve, 2023). That is, much if not most of human learning (both within and outside formal education) occurs because of our interest and curiosity in activities, from which we acquire knowledge and skills. Research in SDT suggests that sustained playful learning involves experiencing a sense of autonomy and competence, which are often richly afforded within game environments (Rigby & Ryan, 2011).

Carefully applying these concepts to the challenge posed by LLMs for expertise development may turn the outlined risks into promising learning opportunities. The idea is the same as the one exemplified by Papert's approach (1980) to utilize computers as educational tools. Instead of seeing LLMs as possibilities to outsource task accomplishment, they are understood as tools that can be utilized to engage in a meaningful activity. The interface, which has been the Logo language in Papert's case (1980), now is, for instance, ChatGPT, allowing to provide prompts that steer the underlying LLM in the desired direction. In this case, the meaningful product, is not necessarily an image, but can be a manuscript, some computer code, or any piece of text. The specific expertise required to be acquired to make LLMs work in such a useful way has become known as *prompt engineering*.

Prompt Engineering as a Form of Playful Interaction with LLMs

Prompt engineering generally refers to the iterative process in which users fine-tune their textual inputs to achieve a desired output from the LLM (Meskó, 2023). It has been recognized as an essential competence within future digital literacy (Eager & Brunton, 2023; Korzynski et al., 2023), eventually enabling to fully harness LLMs' potential to provide personalized learning, unlimited practice opportunities, and interactive engagement with immediate feedback (Heston & Khun, 2023). It has been successfully applied in diverse domains including software development (White et al., 2023), entrepreneurship (Short & Short, 2023), art (Oppenlaender et al., 2023), science (Polak & Morgan, 2023), and healthcare (Wang et al., 2023).

Prompt engineering may involve role play or persona modeling (letting the LLM adopt a specific role such as a domain expert in a certain field; Short & Short, 2023), text format, style or tone (Zhu et al., 2023), length and (coding) language restrictions (Shue et al., 2023), question refinement or alternative approaches requests,

flipped interaction patterns (e.g., requesting questions rather than elaboration from the LLM; White et al., 2023), chain-of-thought-prompting (generating intermediate outputs; e.g., “Take a deep breath and work on this problem step-by-step”; Yang et al., 2023), or emotional prompting (e.g., “This is very important for my career”; Li et al., 2023) among many more possible techniques. Noteworthy, identified functional prompt patterns have been found to be generalizable over many different domains (White et al., 2023).

Although optimizing prompts has been shown to be capable of vastly improving the accuracy of outputs generated by LLMs (Li et al., 2023; Yang et al., 2023), the fact remains that the critical evaluation of resulting outputs still requires domain expertise. Critically reviewing the resulting output is just as important as optimizing the prompts (Shue et al., 2023).

Prompt engineering itself can actually be regarded as an expert skill requiring not only expertise within the domain (for the selection of appropriate keywords and prompt content) but also of prompt modifiers and the training data and system configuration settings of the specific LLMs (Oppenlaender et al., 2023). Becoming proficient in prompt engineering thus has an analogous meaning for a user of an LLM as becoming proficient in the Logo language for Papert’s (1980) students. It not only allows one to make use of LLMs efficiently, but in order for it to work, i.e., to result in reliable and useful output, it entails practicing exactly that domain expertise which it presupposes. Given the necessary expertise, prompt engineering can thus become a form of playful interaction with LLMs, exploring various aspects of a topic by varying prompt patterns and techniques. Under those circumstances, the domain expert’s intrinsic interest in the reliability and usefulness of results produced in cooperation with LLMs might provide some protection from over-reliance on a single output and associated risks of more narrowly directed LLM employments.

However, such risks might be more severe for learners who are not yet domain experts but are presently on their way to developing such expertise. Their primary goals may be less intrinsically motivated but eventually correspond rather to the mere accomplishment of educational tasks like the submission of seminar papers, homework, or sample calculations. In light of the especially low opportunity costs of LLMs, supporting a playful approach for working with them also under those circumstances may require more than to appeal to individual integrity and virtue. Such support, however, may then be accomplished by providing a learning environment in which playing becomes a natural form of activity (Plass et al., 2020) and a designed pathway to learning. That means, such support may be provided by a pedagogy of learning based on games.

Game-Based Learning as a Way to Harness the Full Potential of Human-AI Interaction in Education

Games, in both non-digital and digital forms, have repeatedly proven valuable for learning, training, and education (Dillon et al., 2017; Pahor et al., 2022; Pasqualotto et al., 2022). They provide space for playful learning experiences, allow room for experimentation, and provide safe spaces for graceful failure, a crucial component

for learning with games, allowing the players to learn from mistakes and motivating them to practice until feeling confident (Plass et al., 2015).

Due to their capabilities in capturing and holding people's attention and in fostering sustained engagement and long-term loyalties, games have further become role models for engaging learners (Rigby, 2014) and citizens to solve complex scientific problems (Cooper et al., 2010; Spiers et al., 2023). Well-designed games can indeed promote both the required persistence in activities for practice and high quality of engagement that can foster deep human learning and problem solving (Barz et al., 2023; Hu et al., 2022; Ryan & Rigby, 2020). The extension of SDT (Ryan & Deci, 2000, 2017) based on research on video games (Ryan et al., 2006), technology design (Calvo & Peters, 2014), or digital learning (Sørøbø et al., 2009) has shown in which ways psychological satisfactions for autonomy, competence, and relatedness can be evoked or undermined and thus affect players' intrinsic motivation and sustained engagement (Ryan & Rigby, 2020). In games, a complex set of skills is challenged in a constrained environment in which those skills can be explored, analyzed, manipulated, extended (Ryan & Rigby, 2020), or in other words: practiced. Thereby, ample opportunities allow experiences of autonomy, competence, and relatedness fuelling intrinsic motivation. "In a well-designed game, the learning becomes its own reward" (Ryan & Rigby, 2020, p. 169).

The problem-based gaming model (Kiili, 2007) further emphasizes the meaning of experiential learning and reflection in educational games. It is argued that the ability to reflect may be the main factor determining who learns effectively from experience (Kiili, 2007). This is especially true for games that require problem-solving (e.g., simulation games). In the model, the level of reflection concerns whether the player considers the consequences of their actions and the changes in the game world to create better playing strategies (double-loop learning) or merely applies the previously formed playing strategy (single-loop learning). Games that trigger double-loop learning are effective because they persuade players to test different kinds of hypotheses and consider the learning content deeply from several perspectives. The challenge of educational game design is to design game mechanics that trigger such meaningful reflection practices.

Games as a Culture Medium for the Development of Expertise

Games naturally serve all three learning processes facilitating the development of expertise. By providing ample space for playful engagement, they support the transformation of experiential into conceptual knowledge. By being—in contrast to free-form play—yet structured by explicit rule sets and specific goals (Deterding et al., 2011), they also require and thus facilitate the transformation of conceptual into experiential knowledge. Finally, as outlined above, they invite diverse forms of reflection serving the further development of problem-solving strategies as well as metacognitive and self-regulatory skills.

The capabilities of games to invite reflection are further emphasized by the fact that successful games have repeatedly been identified as sources of spontaneously emergent culture. Affinity groups (Gee, 2005) may emerge (online or offline) in

which players meet to communicate, reflect, and influence game rules, extend new game content, and contribute to game development (Brown, 2016), engage in theorycrafting (Choontanom & Nardi, 2012) and peer-to-peer apprenticeship (Steinkuehler & Oh, 2012). Both the explication of experiential knowledge into conceptual knowledge and reflecting on both knowledge types happen naturally in such spontaneously forming collaborative spaces.

The emergence of those spaces is not induced by top-down mechanisms (e.g., by game developers) but happens horizontally within the game community (Steinkuehler & Tsaasan, 2020). For instance, in the *Just Press Play* project (Decker & Lawley, 2013), investigating the effect of gamification on undergraduate experience in computer science, students spontaneously requested access to computer labs for tutoring other students for free, on their own time and out of their own desire. In addition, a lively community of educators emerged, constantly creating new learning environments and trying to include the game in the class room against all technical and bureaucratic odds. After the release of *Minecraft*, communities emerged, modifying the game and creating content way beyond the games' original intended meaning and functionalities (Nebel et al., 2016). Users—and mostly pupils—used the games' mechanics to create functioning CPUs, landscapes of their favorite books or sustainable environments, all in their free time. Those are both unforeseen and astonishing results. Not only provide they examples of what the notion of “learning outcomes” in game-based learning can actually encompass: the spontaneous emergence of teachers or experts from a community of students or novices (Steinkuehler & Tsaasan, 2020). They also provide examples of what potential game-based learning might bear for education.

Furthermore, they provide examples of how games can foster spontaneous profound engagement with the learning material far beyond a mere accomplishment of tasks. When within well-designed games, in which the basic needs of autonomy, competence, and relatedness are met, learning becomes its own reward (Ryan & Rigby, 2020), the option to outsource cognitive efforts to LLMs becomes less tempting. Instead, well-designed games might even foster the motivation to utilize LLMs for engaging deeper with the content and finding out more. That means, game environments might provide novices with a flavor of that kind of intrinsic interest that may protect domain experts from over-reliance and associated risks.

Yet Where Are All the Educational Games?

However, if games hold such an educational potential, the question needs to be addressed: Why have they not become much more abundant in schools and universities? One simple reason is that making good games, i.e., games that satisfy basic psychological needs (Ryan & Rigby, 2020), is tough. Even established developers in the entertainment game industry, i.e., in the business of manufacturing fun, repeatedly fail to deliver and are regularly hit with closures and layoffs (Hodent, 2018), whereas some of the most successful games started as low-budget side projects. Educational games face many additional challenges.

On a socio-cultural dimension (Fernández-Manjón et al., 2015), an issue is social rejection of games, which may be reduced by improving society's understanding of games as another form of cultural good, and informing stakeholders (students, educators, and parents) about the social potential and positive effects of video games (Granic et al., 2014) and their usefulness in education (Bourgonjon et al., 2010). At the same time, violence, sexism, and discrimination are advised to be avoided in the design of educational games (Fernández-Manjón et al., 2015).

Along an educational dimension, limited accessibility to educational games can prevent their further adoption in education (Fernández-Manjón et al., 2015). Whereas creating and maintaining user manuals and best practice guides are ways to facilitate accessibility (Fernández-Manjón et al., 2015), both require further structural support. The latter can be provided by simultaneous support and creation of communities of practice (Wenger, 1998) allowing participation in development processes (Moreno-Ger et al., 2008) and knowledge production and transfer between educators, developers, and researchers (Fernández-Manjón et al., 2015; Hébert et al., 2021).

Along a technological dimension, limited accessibility to technology is an issue (Hébert et al., 2021). Lowering development costs and developing environments that allow educators some game development without requiring substantial programming skills and specific game development expertise are regarded as necessary steps to address this issue (Fernández-Manjón et al., 2015).

LLMs as an Opportunity for Harnessing the Potential of Games Within Education

In this context, LLMs or more generally generative AI tools have the potential to transform game-based learning practices and—again similarly to the use of computers in Papert's class (1980)—could even become once more part of their own remedy regarding the challenge they pose for education. This, however, warrants a new pedagogy of learning with artificial intelligence.

In particular, we identified two use scenarios in which generative AI tools can boost the use of games in educational settings. First, generative AI tools provide new ways to implement making games for learning approaches (Kafai & Burke, 2015), in which students learn educational content by designing and making games. Second, teachers and educators can utilize AI tools to gamify their learning materials or even create fully-fledged learning games for their students. In the following, we consider how LLMs can be utilized in these scenarios.

Learning by Generating Games

Making games for learning is another prime example of a constructionist learning activity (Kafai & Burke, 2015) similar to Papert's (1980) early use of computers in the classroom discussed above. Kafai and Burke (2015) argue that we are witnessing a paradigmatic shift toward constructionist gaming, in which students design games for learning instead of just consuming games created by professional developers.

We believe that generative AI tools will further accelerate this shift. LLMs have the potential to make game creation more accessible for novices in a similar way as block-based visual programming environments like *Scratch* (Resnick et al., 2009) lowered the demands to program interactive stories and animations in educational settings. The pedagogical idea behind learning by generating games relies mostly on the assumption that game-making activities help students reformulate their understanding of the subject matter (educational content) and express their personal understanding and ideas about the subject (Kafai, 2006). In addition, generating games using AI's technical backup can be open and creative, allowing for experiences of autonomy and competence essential to sustained interest and intrinsic motivation (Ryan & Rigby, 2020). As the technicalities of programming can be largely outsourced to LLMs, students can focus more on the topic and game design aspects.

A recent study indicates that game-designing activities can be even more beneficial, especially for the long-term retention of knowledge, than learning by playing games (Chen & Liu, 2023). Furthermore, Resnick et al., (2009) have emphasized that digital fluency requires more than just interacting with media; it requires an ability to collaboratively design, create, and invent with media. Similar abilities are needed when creating games with the help of LLMs and seem now more important than ever. However, making games with LLMs also imposes unique requirements for students as well as for teachers who are orchestrating the game-making activities.

We coined the term *prompting pedagogy* to capture fundamental pedagogical practices involved in generating games or other digital outputs with the help of LLMs going beyond prompt engineering as discussed above and constituting one aspect of a new pedagogy of learning with AI. While prompt engineering will be a crucial competence for harnessing the potential of AI in education (Eager & Brunton, 2023), we also want to emphasize that the ability to critically evaluate generated outputs and its facilitation by existing (domain) expertise are equally important (Dwivedi et al., 2023; Krügel et al., 2023). This critical evaluation informs the crafting of prompts leading to meaningful and constructive dialogue with LLMs. Such cumulative and continuous dialog is crucial when using LLMs in complex tasks like game-making. Moreover, using LLMs in such a reflective and critical manner enhances critical thinking and problem-based learning (Bernabei et al., 2023; Sallam, 2023; Shue et al., 2023).

It is evident that effective prompting is challenging, and students need support to develop adequate prompting skills to generate games with LLMs. Prompting pedagogy for game-making also involves the preparation of support materials (e.g., prompting templates for different purposes) and sequencing the prompting activities to specific phases (e.g., idea generation, core design, prototyping, and assessment).

Even though the use of LLMs plays a crucial role in the suggested learning by generating games approach, the design and production activities need to be integrated into a meaningful teaching process. For example, the creative thinking spiral process (imagine, create, play, share, reflect, imagine, and so forth) can be adapted to the learning by generating games approach (Resnick, 2009). According to Resnick (2009, p. 1), in this process, “people imagine what they want to do, create a project based on their ideas, play with their creations, share their ideas and creations with others, and reflect on their experiences—all of which leads them to

imagine new ideas and new projects.” This thus provides a way to emphasize play-testing with peers (sharing and testing game prototypes and games) as well as reflective discussion sessions about prompting and game design strategies in LLMs-based game-making projects.

Overall, learning by generating games promotes a creative, experimental, playful, and inclusive learning culture that aims to support the learning of academic content while preparing students for utilizing generative AI tools effectively and creatively in different contexts. As teachers have a significant role in this approach, a starting point for them may be to generate at least one learning game with an LLM before applying the learning approach in their teaching. Such first-hand experience can facilitate perceiving the affordances that LLMs provide, preparing support materials for students, and planning the workflow of activities.

The use of generative AI for developing learning games may also help to decrease the barriers to the creation of low-budget game productions and educational games. One problem with educational games is that we have become more and more accustomed to big-budget releases. Many educational games seem degraded by comparison (e.g., poor graphics and mechanics) and are thus perceived as boring or unappealing. A reasonable utilization of LLMs for game development could eventually help to close this gap. Moreover, since the activities are learner generated, they may well engender a different kind of interest and sense of ownership than studio produced educational outputs.

Gamifying Learning Materials

Generative AI provides many low-threshold possibilities for educators to gamify their teaching or to generate learning games for teaching. That is, LLMs and generative AI might establish themselves as a useful tool for developing (educational) games, for instance, by supporting the generation of artwork, code, or game levels (Nasir & Togelius, 2023; Todd et al., 2023). LLMs can further assist educators in the analysis, design, evaluation, and development phases of game creation projects, allowing, for instance, the adaptation of popular board games such as *Monopoly* for specific learning purposes (Gatti Junior et al., 2023).

It may eventually not matter whether the game makers are students or educators; the generation of games with LLMs requires a playful, experimental, and iterative style of engagement in which game makers continually reassess their goals, explore new solutions, and imagine new possibilities based on the generated outputs and dialogue with LLMs. Resnick and Rosenbaum (Resnick & Rosenbaum, 2013) called such a bottom-up approach “tinkering.” As highlighted above, one of the key skills of the successful generation of games with LLMs is prompting and critical evaluation of generated outputs, which requires expertise such tinkering might enhance.

As game development is usually a highly interdisciplinary process requiring expertise in various areas, LLMs might be used to complement individuals’ skills in a particular area. For instance, it might allow an educator with expertise in the pedagogical approach for a given problem and an idea for the game design to implement a working prototype of an educational game, which would have been significantly more difficult for the educator without using generative AI

technologies. Furthermore, as game design is a very complex activity, it is important to break complex prompts into a series of small, specific steps and phases, starting from the idea generation and identification of instructional approaches and core game mechanics. For example, *chain-of-thought prompting* (generating intermediate outputs) or *role prompting* (giving the LLM, e.g., the specific role of an instructional designer or target group player) can increase the model's contextual tenability and enhance the quality of outputs.

Conclusions and General Remarks

On balance, implementing insights from game-based learning in educational contexts is far from a straightforward task. However, game-based learning research has revealed that well-designed games indeed address, challenge, and promote players holistically, incorporating all cognitive (Mayer, 2020), affective (Loderer et al., 2020), motivational (Ryan & Rigby, 2020), and sociocultural (Steinkuehler & Tsasan, 2020) aspects of the human condition. Applications of game-based learning in science, technology, engineering, and mathematics (Klopfer & Thompson, 2020), or the development of educational games for critical thinking (Butcher et al., 2017) or social problem-solving (Ang et al., 2017) indicate at least the potential games may have for fostering deep engagement with the learning material and continuous practice of expertise. Utilizing LLMs for learning by generating games and purposefully gamifying learning materials may allow educators to fully harness the potential of games toward a new pedagogy of learning with AI.

The potential of playful and game-based learning we see for education is strongly related to games' motivational and engaging power, that "in play, the aim is play itself" (Flanagan, 2009). Even if the activities associated with the playful engagement encountered in games could be delegated to AI support, who would want this—because in this context, it would be outsourcing the fun and intrinsic satisfactions of play. That would be like delegating joy to a robot. Even if we could, why should we want that?

The notion of (good) practice has since Aristotle (2020) involved the aspect of bearing its meaning in itself, a quality which practice has, according to Rousseau and Schiller (Greipl et al., 2020), in common with play. It seems as if the advent of AI challenges us as educators to *remember* and *revive* research and its teaching, as such practice calls for the creation and cultivation of playful spaces within education. While this perspective is certainly not about advocating that we redesign each class into a game promising enjoyment or entertainment, we think that game-based learning could be especially valuable in taking advantage of the educational capabilities of AI, which themselves require capable human partnership.

Funding Open access funding provided by University of Graz. The authors acknowledge the financial support of the University of Graz. Kristian Kiili was supported by the Strategic Research Council (SRC) established within the Research Council of Finland [Grants: 335625, 358250].

Data Availability Not applicable to this article as no datasets were generated or analysed during the current study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abd-alrazaq, A., AlSaad, R., Alhuwail, D., Ahmed, A., Healy, P. M., Latifi, S., Aziz, S., Damseh, R., Alabed Alrazak, S., & Sheikh, J. (2023). Large language models in medical education: Opportunities, challenges, and future directions. *JMIR Medical Education*, 9, e48291. <https://doi.org/10.2196/48291>
- Ali, S. R., Dobbs, T. D., Hutchings, H. A., & Whitaker, I. S. (2023). Using ChatGPT to write patient clinic letters. *The Lancet Digital Health*, 5(4), e179–e181. [https://doi.org/10.1016/S2589-7500\(23\)00048-1](https://doi.org/10.1016/S2589-7500(23)00048-1)
- Ang, R. P., Tan, J. L., Goh, D. H., Huan, V. S., Ooi, Y. P., & Boon, J. S. T. (2017). A game-based approach to teaching social problem-solving skills. In R. Z. Zheng & M. K. Gardner (Eds.), *Handbook of research on serious games for educational applications* (pp. 115–148). IGI Global.
- Aristotle. (2020). *The Nicomachean Ethics* (A. Beresford, Trans.). Penguin Classics. (Original work published ca. 350 B.C.E.)
- Ariyaratne, S., Iyengar, K. P., Nischal, N., ChittiBabu, N., & Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Skeletal Radiology*. <https://doi.org/10.1007/s00256-023-04340-5>
- Barz, N., Benick, M., Dörrenbächer-Ulrich, L., & Perels, F. (2023). The effect of digital game-based learning interventions on cognitive, metacognitive, and affective-motivational learning outcomes in school: A meta-analysis. *Review of Educational Research*, 003465432311677. <https://doi.org/10.3102/00346543231167795>
- Bernabei, M., Colabianchi, S., Falegnami, A., & Costantino, F. (2023). Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence*, 5, 100172. <https://doi.org/10.1016/j.caeai.2023.100172>
- Biswas, S. S. (2023). Potential use of Chat GPT in global warming. *Annals of Biomedical Engineering*. <https://doi.org/10.1007/s10439-023-03171-8>
- Bourgonjon, J., Valcke, M., Soetaert, R., & Schellens, T. (2010). Students' perceptions about the use of video games in the classroom. *Computers and Education*, 54(4), 1145–1156. <https://doi.org/10.1016/j.compedu.2009.10.022>
- Brown, J. K. (2016). *To literacy and beyond: The poetics of Disney Infinity 3.0 as facilitators of new literacy practices (Master's thesis)*. University of California, Irvine.
- Butcher, K. R., Runburg, M., & Altizer, R. (2017). Dino Lab: Designing and developing an educational game for critical thinking. In R. Z. Zheng & M. K. Gardner (Eds.), *Handbook of research on serious games for educational applications* (pp. 115–148). IGI Global.
- Calvo, R. A., & Peters, D. (2014). *Positive computing: Technology for wellbeing and human potential*. MIT Press.
- Campitelli, G., & Gobet, F. (2011). Deliberate practice: Necessary but not sufficient. *Current Directions in Psychological Science*, 20(5), 280–285. <https://doi.org/10.1177/0963721411421922>
- Chen, S., & Liu, Y.-T. (2023). Learning by designing or learning by playing? A comparative study of the effects of game-based learning on learning motivation and on short-term and long-term

- conversational gains. *Interactive Learning Environments*, 31(7), 4309–4323. <https://doi.org/10.1080/10494820.2021.1961159>
- Choontanom, T., & Nardi, B. (2012). Theorycrafting: The art and science of using numbers to interpret the world. In C. Steinkuehler, K. Squire, & S. Barab (Eds.), *Games, Learning, and Society* (1st ed., pp. 185–209). Cambridge University Press. <https://doi.org/10.1017/CBO9781139031127.017>
- Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., & Players, F. (2010). Predicting protein structures with a multiplayer online game. *Nature*, 466(7307), 756–760. <https://doi.org/10.1038/nature09304>
- Decker, A., & Lawley, E. L. (2013). Life’s a game and the game of life: How making a game out of it can change student behavior. *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, 233–238. <https://doi.org/10.1145/2445196.2445269>
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining “gamification.” *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. <https://doi.org/10.1145/2181037.2181040>
- Deutscher Ethikrat. (2023). *Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz. Stellungnahme*. Deutscher Ethikrat. <https://www.ethikrat.org/themen/aktuelle-ethikratthemen/mensch-und-maschine/>. Accessed 4 May 2023.
- Dillon, M. R., Kannan, H., Dean, J. T., Spelke, E. S., & Duflo, E. (2017). Cognitive science in the field: A preschool intervention durably enhances intuitive but not formal mathematics. *Science*, 357(6346), 47–55. <https://doi.org/10.1126/science.aal4724>
- Dunn, C., Hunter, J., Steffes, W., Whitney, Z., Foss, M., Mammino, J., Leavitt, A., Hawkins, S. D., Dane, A., Yungmann, M., & Nathoo, R. (2023). Artificial intelligence–derived dermatology case reports are indistinguishable from those written by humans: A single-blinded observer study. *Journal of the American Academy of Dermatology*, S019096222300587X. <https://doi.org/10.1016/j.jaad.2023.04.005>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching and Learning Practice*, 20(5). <https://doi.org/10.53761/1.20.5.02>
- Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature*, 613(7944), 423–423. <https://doi.org/10.1038/d41586-023-00056-7>
- Elvira, Q., Imants, J., Dankbaar, B., & Segers, M. (2017). Designing education for professional expertise development. *Scandinavian Journal of Educational Research*, 61(2), 187–204. <https://doi.org/10.1080/00313831.2015.1119729>
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100(3), 363–406. <https://doi.org/10.1037/0033-295X.100.3.363>
- Fernández-Manjón, B., Moreno-Ger, P., Martínez-Ortiz, I., & Freire, M. (2015). Challenges of serious games. *EAI Endorsed Transactions on Game-Based Learning*, 2(6), 150611. <https://doi.org/10.4108/eai.5-11-2015.150611>
- Flanagan, M. (2009). *Critical play: Radical game design*. MIT Press.
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Guo, Q., Wang, M., & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. <https://doi.org/10.48550/ARXIV.2312.10997>
- Gatti Junior, W., Marasco, E., Kim, B., Behjat, L., & Eggermont, M. (2023). How ChatGPT can inspire and improve serious board game design. *International Journal of Serious Games*, 10(4), 33–54. <https://doi.org/10.17083/ijsg.v10i4.645>
- Gee, J. P. (2005). Semiotic social spaces and affinity spaces: From *The Age of Mythology* to today’s schools. In D. Barton & K. Tusting (Eds.), *Beyond communities of practice* (1st ed., pp. 214–232). Cambridge University Press. <https://doi.org/10.1017/CBO9780511610554.012>
- Google. (2023). *Bard [large language model]*. <https://bard.google.com/>. Accessed 6 Dec 2023.

- Granic, I., Lobel, A., & Engels, R. C. M. E. (2014). The benefits of playing video games. *American Psychologist*, 69(1), 66–78. <https://doi.org/10.1037/a0034857>
- Greipl, S., Moeller, K., & Ninaus, M. (2020). Potential and limits of game-based learning. *International Journal of Technology Enhanced Learning*, 12(4), 363. <https://doi.org/10.1504/IJTEL.2020.110047>
- Hambrick, D. Z., Macnamara, B. N., Campitelli, G., Ullén, F., & Mosing, M. A. (2016). Beyond born versus made. In *Psychology of learning and motivation* (Vol. 64, pp. 1–55). Elsevier. <https://doi.org/10.1016/bs.plm.2015.09.001>
- Hambrick, D. Z., Oswald, F. L., Altmann, E. M., Meinz, E. J., Gobet, F., & Campitelli, G. (2014). Deliberate practice: Is that all it takes to become an expert? *Intelligence*, 45, 34–45. <https://doi.org/10.1016/j.intell.2013.04.001>
- Hébert, C., Jenson, J., & Terzopoulos, T. (2021). “Access to technology is the major challenge”: Teacher perspectives on barriers to DGBL in K-12 classrooms. *E-Learning and Digital Media*, 18(3), 307–324. <https://doi.org/10.1177/2042753021995315>
- Heston, T. F., & Khun, C. (2023). Prompt engineering in medical education. *International Medical Education*, 2(3), 198–205. <https://doi.org/10.3390/ime2030019>
- Hodent, C. (2018). *The gamer’s brain*. CRC Press.
- Homer, B. D., Raffaele, C., & Henderson, H. (2020). Games as playful learning: Implications of developmental theory for game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 25–52). MIT Press.
- Hosseini, M., Gao, C. A., Liebovitz, D. M., Carvalho, A. M., Ahmad, F. S., Luo, Y., MacDonald, N., Holmes, K. L., & Kho, A. (2023). An exploratory survey about using ChatGPT in education, healthcare, and research [Preprint]. *Medical Ethics*. <https://doi.org/10.1101/2023.03.31.23287979>
- Howard, A., Hope, W., & Gerada, A. (2023). ChatGPT and antimicrobial advice: The end of the consulting infection doctor? *The Lancet Infectious Diseases*, 23(4), 405–406. [https://doi.org/10.1016/S1473-3099\(23\)00113-5](https://doi.org/10.1016/S1473-3099(23)00113-5)
- Hu, Y., Gallagher, T., Wouters, P., Van Der Schaaf, M., & Kester, L. (2022). Game-based learning has good chemistry with chemistry education: A three-level meta-analysis. *Journal of Research in Science Teaching*, 59(9), 1499–1543. <https://doi.org/10.1002/tea.21765>
- Hutson, M. (2022). Could AI help you to write your next paper? *Nature*, 611(7934), 192–193. <https://doi.org/10.1038/d41586-022-03479-w>
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873–15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Johnson, A. (2023). ChatGPT in schools: Here’s where it’s banned—And how it could potentially help students. *Forbes*. <https://www.forbes.com/sites/ariannajohnson/2023/01/18/chatgpt-in-schools-heres-where-its-banned-and-how-it-could-potentially-help-students>. Accessed 6 Dec 2023.
- Kafai, Y. B. (2006). Playing and making games for learning: Instructionist and constructionist perspectives for game studies. *Games and Culture*, 1(1), 36–40. <https://doi.org/10.1177/1555412005281767>
- Kafai, Y. B., & Burke, Q. (2015). Constructionist gaming: Understanding the benefits of making games for learning. *Educational Psychologist*, 50(4), 313–334. <https://doi.org/10.1080/00461520.2015.1124022>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kiili, K. (2007). Foundation for problem-based gaming. *British Journal of Educational Technology*, 38(3), 394–404. <https://doi.org/10.1111/j.1467-8535.2007.00704.x>
- Klopfer, E., & Thompson, M. (2020). Game-based learning in science, technology, engineering, and mathematics. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 387–408). MIT Press.
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Korzynski, P., Mazurek, G., Krzypkowska, P., & Kurasinski, A. (2023). Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT. *Entrepreneurial Business and Economics Review*, 11(3), 25–37. <https://doi.org/10.15678/EBER.2023.110302>

- Krügel, S., Ostermaier, A., & Uhl, M. (2023). ChatGPT's inconsistent moral advice influences users' judgment. *Scientific Reports*, 13(1), 4569. <https://doi.org/10.1038/s41598-023-31341-0>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36(6), 661–679. <https://doi.org/10.1017/S0140525X12003196>
- Li, C., Wang, J., Zhang, Y., Zhu, K., Hou, W., Lian, J., Luo, F., Yang, Q., & Xie, X. (2023). Large language models understand and can be enhanced by emotional stimuli. <https://doi.org/10.48550/ARXIV.2307.11760>
- Loderer, K., Pekrun, R., & Plass, J. L. (2020). Emotional foundations of game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 111–151). MIT Press.
- Mayer, R. E. (2020). Cognitive foundations of game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 83–110). MIT Press.
- Meskó, B. (2023). Prompt engineering as an important emerging skill for medical professionals: Tutorial. *Journal of Medical Internet Research*, 25, e50638. <https://doi.org/10.2196/50638>
- Microsoft. (2023). Bing Chat [large language model]. <https://www.bing.com/chat>. Accessed 6 Dec 2023.
- Miller, E. M., Walton, G. M., Dweck, C. S., Job, V., Trzesniewski, K. H., & McClure, S. M. (2012). Theories of willpower affect sustained learning. *PLoS One*, 7(6), e38680. <https://doi.org/10.1371/journal.pone.0038680>
- Mogali, S. R. (2023). Initial impressions of CHATGPT for anatomy education. *Anatomical Sciences Education*, ase.2261. <https://doi.org/10.1002/ase.2261>
- Molenaar, I. (2021). Personalisation of learning: Towards hybrid human-AI learning technologies. In *OECD digital education outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD Publishing. <https://read.oecd.org/10.1787/2cc25e37-en?format=html>. Accessed 26 Jun 2023.
- Moreno-Ger, P., Martínez-Ortiz, I., Sierra, J. L., & Fernández-Manjón, B. (2008). A content-centric development process model. *Computer*, 41(3), 24–30. <https://doi.org/10.1109/MC.2008.73>
- Morris, M. R., Sohl-dickstein, J., Fiedel, N., Warkentin, T., Dafoe, A., Faust, A., Farabet, C., & Legg, S. (2023). Levels of AGI: Operationalizing progress on the path to AGI. <https://doi.org/10.48550/ARXIV.2311.02462>
- Nasir, M. U., & Togelius, J. (2023). Practical PCG through large language models. <https://doi.org/10.48550/ARXIV.2305.18243>
- Nebel, S., Schneider, S., & Rey, G. D. (2016). Mining learning and crafting scientific experiments: A literature review on the use of minecraft in education and research. *Journal of Educational Technology and Society*, 19(2), 355–366.
- Ninaus, M., & Sailer, M. (2022). Closing the loop – The human role in artificial intelligence for education. *Frontiers in Psychology*, 13, 956798. <https://doi.org/10.3389/fpsyg.2022.956798>
- OpenAI. (2023). ChatGPT [large language model]. <https://chat.openai.com/chat>. Accessed 6 Dec 2023.
- Oppenlaender, J., Linder, R., & Silvennoinen, J. (2023). Prompting AI art: An investigation into the creative skill of prompt engineering. <https://doi.org/10.48550/ARXIV.2303.13534>
- Pahor, A., Seitz, A. R., & Jaeggi, S. M. (2022). Near transfer to an unrelated N-back task mediates the effect of N-back working memory training on matrix reasoning. *Nature Human Behaviour*, 6(9), 1243–1256. <https://doi.org/10.1038/s41562-022-01384-w>
- Papert, S. (1980). *Mindstorms. Children, computers and powerful ideas*. Basic Books.
- Pasqualotto, A., Altarelli, I., De Angeli, A., Menestrina, Z., Bavelier, D., & Venuti, P. (2022). Enhancing reading skills through a video game mixing action mechanics and cognitive training. *Nature Human Behaviour*, 6(4), 545–554. <https://doi.org/10.1038/s41562-021-01254-x>
- Patterson, R. E., Pierce, B. J., Bell, H. H., & Klein, G. (2010). Implicit learning, tacit knowledge, expertise development, and naturalistic decision making. *Journal of Cognitive Engineering and Decision Making*, 4(4), 289–303. <https://doi.org/10.1177/155534341000400403>
- Piaget, J. (1962). *Play, dreams, and imitation in childhood*. Norton.
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Plass, J. L., Homer, B. D., Mayer, R. E., & Kinzer, C. K. (2020). Theoretical foundations of game-based and playful learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 3–24). MIT Press.
- Polak, M. P., & Morgan, D. (2023). Extracting accurate materials data from research papers with conversational language models and prompt engineering. <https://doi.org/10.48550/ARXIV.2303.05352>

- Reeve, J. (2023). Cognitive evaluation theory: The seedling that keeps self-determination theory growing. In R. M. Ryan (Ed.), *The Oxford handbook of self-determination theory* (1st ed., pp. 33–C2P117). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197600047.013.3>
- Resnick, M. (2009, April 4). Sowing the seeds for a more creative society. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09: CHI Conference on Human Factors in Computing Systems, Boston MA USA. <https://doi.org/10.1145/1518701.2167142>
- Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: Programming for all. *Communications of the ACM*, 52(11), 60–67. <https://doi.org/10.1145/1592761.1592779>
- Resnick, M., & Rosenbaum, E. (2013). Designing for tinkering. In M. Honey (Ed.), *Design, make, play: Growing the next generation of STEM innovators* (pp. 163–181). Routledge.
- Rigby, C. S. (2014). Gamification and motivation. In S. P. Walz & S. Deterding (Eds.), *The gameful world* (pp. 113–138). MIT Press.
- Rigby, C. S., & Ryan, R. M. (2011). *Glued to games: How video games draw us in and hold us spellbound*. Praeger.
- Rosenzweig-Ziff, D. (2023). New York City blocks use of the ChatGPT bot in its schools. *The Washington Post*. <https://www.washingtonpost.com/education/2023/01/05/nyc-schools-ban-chatgpt/>. Accessed 6 Dec 2023.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.
- Ryan, R. M., & Rigby, C. S. (2020). Motivational foundations of game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 153–176). MIT Press.
- Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion*, 30(4), 344–360. <https://doi.org/10.1007/s11031-006-9051-8>
- Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare*, 11(6), 887. <https://doi.org/10.3390/healthcare11060887>
- Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, 20(3), 2–9. <https://doi.org/10.3102/0013189X020003002>
- Salvagno, M., Taccone, F. S., & Gerli, A. G. (2023). Can artificial intelligence help for scientific writing? *Critical Care*, 27(1), 75. <https://doi.org/10.1186/s13054-023-04380-2>
- Sarsa, S., Denny, P., Hellas, A., & Leinonen, J. (2022). Automatic generation of programming exercises and code explanations using large language models. *Proceedings of the 2022 ACM Conference on International Computing Education Research - Volume, 1*, 27–43. <https://doi.org/10.1145/3501385.3543957>
- Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights*, 19, e00388. <https://doi.org/10.1016/j.jbvi.2023.e00388>
- Shue, E., Liu, L., Li, B., Feng, Z., Li, X., & Hu, G. (2023). Empowering beginners in bioinformatics with ChatGPT [Preprint]. *Bioinformatics*. <https://doi.org/10.1101/2023.03.07.531414>
- Sørøbø, Ø., Halvari, H., Gulli, V. F., & Kristiansen, R. (2009). The role of self-determination theory in explaining teachers' motivation to continue to use e-learning technology. *Computers and Education*, 53(4), 1177–1187. <https://doi.org/10.1016/j.compedu.2009.06.001>
- Spiers, H. J., Coutrot, A., & Hornberger, M. (2023). Explaining world-wide variation in navigation ability from millions of people: Citizen science project sea hero quest. *Topics in Cognitive Science*, 15(1), 120–138. <https://doi.org/10.1111/tops.12590>
- Steinkuehler, C., & Oh, Y. (2012). Apprenticeship in massively multiplayer online games. In C. Steinkuehler, K. Squire, & S. Barab (Eds.), *Games, learning, and society: Learning and meaning in the digital age* (pp. 185–209). Cambridge University Press. <https://doi.org/10.1017/CBO9781139031127.017>
- Steinkuehler, C., & Tsaasan, A. M. (2020). Sociocultural foundations of game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 177–206). MIT Press.
- Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove. *Nature*, 613(7945), 620–621. <https://doi.org/10.1038/d41586-023-00107-z>

- Todd, G., Earle, S., Nasir, M. U., Green, M. C., & Togelius, J. (2023). Level Generation Through Large Language Models. *Proceedings of the 18th International Conference on the Foundations of Digital Games*, 1–8. <https://doi.org/10.1145/3582437.3587211>
- Tynjälä, P. (2008). Perspectives into learning at the workplace. *Educational Research Review*, 3(2), 130–154. <https://doi.org/10.1016/j.edurev.2007.12.001>
- Tynjälä, P., Slotte, V., Nieminen, J., Lonka, K., & Olkinuora, E. (2006). From university to working life: Graduates' workplace skills in practice. In P. Tynjälä, J. Välimaa, & G. Boulton-Lewis (Eds.), *Higher education and working life: Collaborations, confrontations and challenges* (pp. 77–88). Elsevier Earli.
- Van Dis, E. A. M., Bollen, J., Zuidema, W., Van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224–226. <https://doi.org/10.1038/d41586-023-00288-7>
- Vygotsky, L. S. (1967). Play and its role in the mental development of the child. *Soviet Psychology*, 5(3), 6–18. <https://doi.org/10.2753/RPO1061-040505036>
- Wang, J., Shi, E., Yu, S., Wu, Z., Ma, C., Dai, H., Yang, Q., Kang, Y., Wu, J., Hu, H., Yue, C., Zhang, H., Liu, Y., Li, X., Ge, B., Zhu, D., Yuan, Y., Shen, D., Liu, T., & Zhang, S. (2023). Prompt engineering for healthcare: Methodologies and applications. <https://doi.org/10.48550/ARXIV.2304.14670>
- Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. <https://doi.org/10.48550/ARXIV.2302.11382>
- Wolfram, S. (2023). *What is ChatGPT doing and why does it work?* Wolfram Media.
- Yang, C., Wang, X., Lu, Y., Liu, H., Le, Q. V., Zhou, D., & Chen, X. (2023). Large language models as optimizers. <https://doi.org/10.48550/ARXIV.2309.03409>
- Zhu, J.-J., Jiang, J., Yang, M., & Ren, Z. J. (2023). ChatGPT and environmental research. *Environmental Science & Technology*, acs.est.3c01818. <https://doi.org/10.1021/acs.est.3c01818>
- Zimmerman, A. (2023). A ghostwriter for the masses: ChatGPT and the future of writing. *Annals of Surgical Oncology*, s10434–023–13436–0. <https://doi.org/10.1245/s10434-023-13436-0>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Stefan E. Huber¹  · Kristian Kiili²  · Steve Nebel³  · Richard M. Ryan^{4,5}  · Michael Sailer⁶  · Manuel Ninaus^{1,7} 

✉ Manuel Ninaus
manuel.ninaus@uni-graz.at

¹ Department of Psychology, University of Graz, Graz, Austria

² Faculty of Education and Culture, Tampere University, Tampere, Finland

³ Media Education, Department of Educational Research, University of Potsdam, Potsdam, Germany

⁴ Institute for Positive Psychology and Education, Australian Catholic University, Sydney, NSW, Australia

⁵ College of Education, Ewha Womans University, Seoul, South Korea

⁶ Learning Analytics and Educational Data Mining, University of Augsburg, Augsburg, Germany

⁷ LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany