

Article type

Full paper, double-blind peer review.

Publication history

Received: 13 February 2025. Revised: 29 March 2025.
Accepted: 07 April 2025. Online: 12 May 2024.

Cover image

Badly Disguised Bligh
via flickr.



Special issue Generative AI and education | **More at** <https://doi.org/10.21428/8c225f6e.0390b853>

Large Language Model Chatbots in education: Exploring literature insights on their impact and influence on learning behaviours

Megan Ennion and Ros McLellan

Faculty of Education, University of Cambridge, Cambridge, United Kingdom

Keywords

ChatGPT in education; learning behaviours; LLM chatbots in education; Social Cognitive Theory; Self-determination Theory

Citation

Ennion, M., & McLellan, R. (in press). Large Language Model Chatbots in education: Exploring literature insights on their impact and influence on learning behaviours. **Studies in Technology Enhanced Learning**, 4(1), 1-23

<https://doi.org/10.21428/8c225f6e.540b41b5>

Abstract

Since the launch of ChatGPT-3.5 in 2022, Large Language Model (LLM) chatbots have garnered significant attention in educational contexts, sparking debates about their potential benefits and drawbacks. While existing research primarily focuses on academic outcomes, there remains a notable gap in understanding how these tools influence key learning behaviours such as resilience, effort, perseverance, and challenge-seeking. This paper begins by introducing LLM chatbots and providing an overview of current literature on their impact in education, highlighting both their advantages and limitations. Underscoring a gap in research, this paper draws attention to the overlooked potential of these chatbots to influence learning behaviours which are valuable for academic and lifelong success. By examining emerging research on LLM chatbot capabilities and their impact on the learning experience through the lens of established psychological theories, such as Social Cognitive Theory and Self-Determination Theory, this paper discusses the potential influence of LLM chatbots on key learning behaviours, alongside practical and psychological challenges that may hinder their effectiveness. Bringing attention to this important, yet under-researched relationship, this paper advocates for empirical studies to explore the relationship between LLM chatbots and learning behaviours, as this understanding is vital for the ethical and effective integration of these technologies in educational settings.

1. Introduction

When ChatGPT-3.5 was released in November 2022, its advanced and human-like capabilities quickly captured widespread attention. In education, interest in the role of Large Language Model (LLM) chatbots, such as ChatGPT, grew rapidly, with some praising their transformative potential while others warned of existential risks to educational practices. Over the past two years, as educators and researchers have worked to understand these technologies' applications and implications - often without clear insight into their underlying mechanics - developers have continued to release and market increasingly sophisticated models and educational technology tools. Meanwhile, students worldwide have increasingly turned to these tools, for better or worse, with varying levels of guidance.

During this period, a growing body of research has emerged examining various dimensions of LLM chatbots in education and producing mixed findings that reflect the complex and nuanced nature of their impact. Amid these conflicting results and heated debates, significant gaps in research and understanding persist due to the field's infancy. One critical but often overlooked consideration is the impact of these tools on students' learning behaviours, which are widely regarded as instrumental for success both within education and throughout life.

This review explores the impact of Large Language Model (LLM) chatbots within educational contexts, examining their effect through the lens of foundational theory, such as Social Cognitive Theory and Self-Determination Theory, alongside emerging research and literature in the field. Following discussion of the methodological approach employed, Section 3 provides an overview of research on LLM chatbots in education, highlighting potential and limitations. From this, it identifies an important gap in this research, positing that one potential benefit of LLM chatbots has been overlooked: their potential to influence learning behaviours. Focusing on four key mastery learning behaviours - resilience, effort, perseverance, and challenge-seeking - Section 4 underscores the important role these behaviours play in supporting greater success. In the absence of empirical studies explicitly testing this relationship, the following sub-sections draw on existing research and theory to conceptually explore the unexamined effects of LLM chatbots on student learning behaviours. Section 5 examines the potential of LLM chatbots to promote these behaviours, while the practical limitations which may prevent this potential from being realised are discussed in Section 6. Building on these discussions, this review

concludes that examining the impact of LLM chatbots on learning behaviours is a valuable and necessary area of research with significant potential and calls for empirical research into this relationship.

2. Methodological approach

2.1 Objective and scope

The objective of this paper was to synthesise insights into the impact of LLM chatbots within education and, in identifying a significant empirical gap, explore their theoretical effects on learning behaviours such as effort, resilience, perseverance, and challenge-seeking. Through this examination, the aim is to propose new directions for empirical research to deepen understanding of these dynamics.

2.2 Source identification and selection criteria

Given the limited empirical studies on the behavioural impacts of LLM chatbots, this paper adopts a conceptual approach, utilising studies that inform the theoretical analysis. Sources were identified through academic databases like ERIC and the British Education Index, as well as seminal works recognised for their contributions to the discourse on LLM chatbots. Over 60 articles were selected, primarily from peer-reviewed journals, ensuring a mix of qualitative and quantitative empirical studies, alongside comprehensive review articles. Priority was given in the selection of empirical studies to those with robust experimental designs or substantial sample sizes. Additionally, articles were selected for their unique contributions to the ongoing conversation about the educational impacts of LLM chatbots.

2.3 Theoretical framework

The analysis is rooted in established theories of learning psychology and behaviour, framing the potential impacts of LLM chatbots within recognised educational and psychological paradigms. This paper integrates these theories with current emerging research to hypothesise how LLM chatbots could potentially enhance or inhibit specific learning behaviours. It links factors such as increased motivation and self-efficacy, associated with positive learning outcomes, to the capabilities of LLM chatbots identified in emerging research. This approach enables use of deductive reasoning to hypothesise the potential influence of chatbots on learning behaviours based on their alignment with these key factors.

In practice, this paper reviews foundational theories that highlight critical behaviours for successful learning and assesses how LLM chatbots support these behaviours through features like instant feedback and personalised interactions. This is accompanied by an analysis of empirical evidence, integrating recent studies that demonstrate chatbots' roles in enhancing or undermining learning factors such as engagement. These data help to strengthen or challenge the theoretical predictions. Finally, through deductive argumentation, chatbots' capabilities are logically connected to theoretically and empirically supported outcomes, hypothesising that LLM chatbots are likely to influence learning behaviours.

2.4 Balanced analysis

The methodology provides a balanced analysis to deepen the exploration of the theoretical arguments. Both the positive and negative impacts of LLM chatbots in education are detailed, addressing counter-arguments by recognising challenges such as dependency risks and the promotion of superficial over deep learning. The varied findings in the literature are explored, highlighting and analysing contradictory evidence and theorising potential reasons for these discrepancies. Despite noting chatbots' potential to positively influence behaviours, specific challenges to this potential are also rigorously challenged, such as the lack of human-like interaction and inherent biases, which could undermine their effectiveness.

2.5 Limitations

As a conceptual paper, this study lays the groundwork for future empirical exploration but does not provide empirical data to conclusively determine the impact of LLM chatbots on learning behaviours. Additionally, while striving for a balanced view by including a broad spectrum of studies, potential biases in study selection are acknowledged.

3. Unpacking the promise and pitfalls of LLM Chatbots in education

3.1 The evolution and challenges of AI Chatbots in education

Chatbots are software programs designed to mimic human conversation (Yang et al., 2023). They use machine learning algorithms to analyse input and generate responses which enable them to engage in conversations and tailor interactions to individual users. Although chatbots have

been used for various purposes across domains since the early 2000s (Labadze, Grigolia & Machaidze, 2023), their adoption in education has been slower due to historically limited capabilities.

Over the past decade, however, advancements in AI have supported the evolution of educational AI chatbots (Kuhail et al., 2023). Around the mid-2010s, Generative AI, which creates new content based on learned patterns, and Adaptive AI, which continuously learns and evolves from interactions, began enhancing the ability of AI chatbots to deliver personalised learning experiences (Kuhail et al., 2023). Notably, in the late 2010s, significant improvements to Large Language Models (LLMs) propelled chatbot capacities and educational potential (Kasneci et al., 2023).

Large Language Models (LLMs), such as the 'GPT' that powers ChatGPT, are a type of AI that employs Machine Learning and Deep Learning to process and generate human-like text from extensive datasets. These models utilise advanced learning techniques to predict the most likely next word or "token" based on patterns and probabilities derived from the vast amounts of text they have been trained on. As a result, chatbots leveraging LLM technology are highly anthropomorphic (Zhu et al., 2023) and can engage in interactions that feel remarkably authentic and human-like (Kohnke, Moorhouse & Zou, 2023; Tlili et al., 2023).

These advances make LLM chatbot technology particularly impressive and unique; however, they also bring with them significant challenges, especially regarding their role in education. The complexity of LLMs' inner workings means it is not yet possible to fully understand or trace the reasoning behind LLM outputs; essentially, these tools function as "black boxes" (Hassija et al., 2023). This lack of transparency makes it difficult to trust these tools and limits the ability to correct errors or address biases, hindering educational value. Additionally, LLM architecture makes them particularly prone to "hallucinations", instances where they appear to fabricate information (Dahlkemper, Lahme & Klein, 2023; Khurma, Albahti, Ali, & Bustanji, 2024). While referred to as "hallucinations", this term can be misleading. LLMs do not make mistakes in the traditional sense; instead, as they merely generate text by predicting the most probable next word without genuine understanding, this results in a mix of accurate and inaccurate information. Consequently, hallucinations are an inherent feature of LLMs and are unlikely to be completely eliminated when relying solely on this technology.

Moreover, building and operating LLMs requires substantial financial, human, and energy resources, raising sustainability concerns and leading to a monopoly in development by a few powerful technology companies based in the United States (US). While the recent emergence of open-source model DeepSeek, which claims to be more energy efficient and environmentally friendly, may address these concerns, claims around the environmental impact remain to be proven. Additionally, as Holmes (2023) and Mintz, Holmes, Liu, and Perez-Ortiz (2023) highlight, the intertwining of private companies and education raises other significant ethical issues, including the stealth commercialisation of education with potential conflicts of interest (Kikerpill & Siibak, 2023), lack of democratic oversight, concerns about unethical data use (Khan, 2024), siloed research (Luckin & Holmes, 2016), and unequal access (Luo et al., 2024). Additionally, as Mintz et al. (2023) observe, there is a worrying lack of research on the effects of these tools on human development and cognition, alongside insufficient large-scale, independent evaluations of their impact within educational systems.

Independent research is crucial to assessing the long-term, widescale impact of LLMs in education. However, as Novianti (2025) highlights, implementing AI in mainstream educational settings involves significant financial, technical, ethical, and training-related challenges (see also Ng et al., 2023, and Walter, 2024, for discussion of skills and training barriers). Additionally, there are significant and documented risks involved in educational institutions partnering with private educational technology companies to deliver these tools. For example, in 2024, the Los Angeles Unified School District (LAUSD) partnered with AllHere Education - named one of Time magazine's "World's Top Edtech Companies of 2024" - to develop an educational chatbot. However, not only did AllHere suddenly collapse, compromising the future of the tool and money invested by LAUSD, but serious concerns were also raised about the handling and fate of the student data it collected (Young, 2024). Similarly, Byju, an educational technology startup that was valued at \$22 billion in March 2022, was in October 2024 declared to be worth "zero" (Singh, 2024). These cautionary tales highlight the instability and ethical risks associated with reliance on these private companies.

Nevertheless, over time, educational LLM chatbots are likely to become more accessible and affordable, increasing their use in education and the need to explore their impact. As UNESCO (Miao, Holmes, Huang, & Zhang, 2021) highlights, in order to critically leverage and integrate these tools into education, it is crucial to understand all the potential benefits and risks.

3.2 Do LLM Chatbots have transformative potential in education?

Given challenges, it is crucial to use a cautious and critical approach to the integration of LLM technology in education. However, educators would be unwise to ignore these tools altogether, as reports show that many students are already using LLMs - often with more positive perceptions and confidence than their teachers - frequently without school knowledge or support (Impact Research, 2024). Additionally, as comprehensive analyses by Kasneci et al. (2023) and Zhu et al. (2023) highlight, LLM chatbots may offer unique educational advantages afforded by their advanced technology, a potential argued to be so significant, it has been described as 'transformative' (Antony & Ramnath, 2023; Hasanein & Sobaih, 2023; Watson, 2025).

Indeed, various studies on LLM chatbots reveal several valuable benefits for education. For instance, a literature review by Labadze, Grigolia, and Machaidze (2023) examined 67 studies and found significant advantages for teachers, including improved pedagogy and time saving, findings consistent with qualitative perceptions reported by ElSayary (2023). Dai, Lai, Lim, and Liu (2023) found that Australian postgraduate students believed ChatGPT enhanced the quality and pace of their work. Assessing the pros and cons of LLM chatbots in education, Khurma et al. (2024) concluded that they can increase student motivation and engagement. Additionally, studies have examined the impact of LLM chatbots on student learning and achievement, providing evidence of their positive effects. For instance, Sun et al. (2024) conducted a mixed-method study on Python support and found that personalised feedback from LLM chatbots enhanced learning. Similarly, Alneyadi and Wardat (2023) conducted a mixed-method experiment with students in the United Arab Emirates and discovered that those using ChatGPT for schoolwork over two weeks achieved higher academic results than the control group without access to the tool. Indeed, these findings are perhaps best captured by a recent systematic review and meta-analysis of experimental studies by Deng et al. (2024), which concluded that ChatGPT improves academic performance, affective-motivational states, and higher-order thinking propensities.

For a more detailed summary of these studies and their key findings, see Table 1.

Table 1. Summary of Studies on the Transformative Potential of LLM Chatbots in Education

Author(s)	Year	Study Focus	Key Findings
Impact Research	2024	Survey of K-12 teachers, students ages 12-18 years, undergraduates, and parents of K-12 students in the US	Found many students reported using LLMs without school knowledge or support, often having more positive perceptions and confidence than teachers.
Kasneji et al.	2023	Commentary on research and uses of LLMs in education	Highlight the unique educational advantages and challenges of LLM chatbots.
Zhu et al.	2023	Theoretical SWOT analysis of ChatGPT in education	Supports the transformative potential of LLM chatbots and proposes methods for integration.
Antony & Ramnath	2023	Qualitative exploration of the impact of AI chatbots on student engagement and support in higher education	Conclude AI chatbots have transformative potential by supporting effective communication, increasing engagement, and providing timely support.
Hasanein & Sobaih	2023	Qualitative interviews exploring ChatGPT in education	Highlight the perceived transformative potential of ChatGPT alongside six positive and negative consequences.
Labadze, Grigolia, & Machaidze	2023	Literature review on LLM chatbots	Examined 67 studies and found significant advantages for teachers, including improved pedagogy and time saving.
ElSayary	2023	Qualitative perceptions of LLM chatbots by teachers	Teachers reported benefits mostly in lesson planning, teaching, and learning, but less in assessment and feedback.
Dai, Lai, Lim, & Liu	2023	Perceptions of ChatGPT by Australian postgraduate students	Found that students believed ChatGPT enhanced the quality and pace of their work.
Khurma et al.	2024	Systematic review using PRISMA methodology of 16 relevant peer-reviewed articles	Concluded that LLM chatbots can increase student motivation and engagement but may lack empathy and contextual understanding.
Sun et al.	2024	Experimental study into LLM chatbots	Observed improved learning in tasks supported by ChatGPT, while overall performance was not better than control groups.
Alneyadi & Wardat	2023	Experimental study into the use of ChatGPT in schoolwork	Found that students using ChatGPT for two weeks achieved higher academic results than those in the control group.
Deng et al.	2024	Systematic review and meta-analysis on ChatGPT	Concluded that ChatGPT improves academic performance, affective-motivational states, and higher-order thinking propensities.

Note on acronyms: LLM = Large Language Model; SWOT = Strengths, Weaknesses, Opportunities, Threats; PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

3.3 Are LLM Chatbots enhancing learning or just creating dependency?

However, despite positive findings, the impact of LLM chatbots appears not to be clear-cut. For example, challenging basic assumptions, Holmes (2023) argues that these tools merely displace, rather than save, teachers' time. Similarly, while some students report benefits from chatbots, these positive reports do not necessarily translate into improved learning outcomes. For instance, research by Kreijkes et al. (2025) found that while students in their study reported a more positive experience with the LLM chatbot compared to traditional note-taking, its support was actually less effective for information retention. In this study, students retained more information from two texts after three days when they used note-taking techniques to aid retention instead of an LLM chatbot, suggesting that LLM chatbot support can be less effective for retention than traditional methods.

Indeed, the debate on the impact of LLM chatbots on learning is marked by contradictory findings, with some studies indicating that learning is detrimentally affected by these tools due to over-reliance. For example, Bastani et al. (2024) compared the effects of LLM chatbot support with no support among Turkish high school students during a 90-minute session. They found that while the groups receiving LLM support showed significant performance gains during practice, their performance declined to levels lower than those of students who had never used the chatbot once it was removed. Similar results were reported by Darvishi et al. (2024). These studies raise questions about the effectiveness of LLM chatbots to support learning, concluding that, while

chatbot support may enhance performance in its presence, it can lead to student dependence, resulting in "distributed metacognition" (Darvishi et al., 2024, p.10) - a division of cognitive tasks between the student and the AI which means that once the AI is removed, performance suffers. Interestingly, Bastani et al. (2024) found that prompting that same LLM chatbot to function more like a tutor, for example, by giving hints rather than complete answers, mitigated the negative effect on performance when removed, suggesting that ensuring cognitive engagement could be an important factor to positive impacts.

For an overview of these studies, refer to Table 2.

3.4 Why do findings on Chatbot learning vary?

Learning findings may vary for several reasons. Firstly, drawing conclusive and meaningful insights on LLM chatbots presents unique challenges. For example, the impact of an LLM chatbot on student learning is likely to depend on the specific chatbot used - including its technological sophistication, ability to tailor responses to individual users, and effectiveness relating to subject content and pedagogy. For instance, Phung et al. (2023) noted that in their assessment of LLM chatbots for providing computer programming support, ChatGPT-4 significantly outperformed ChatGPT-3.5, nearly matching the impact of a human tutor. This observation underscores the importance of technological advancements; however, adjustments in chatbot design and functionality also lead to substantial differences in educational outcomes. For example, studies such as Bastani et al. (2024) suggest that programming LLM chatbots to function as tutors can lead to more positive outcomes than

Table 2. Studies Indicating Limited Learning Enhancement and Increased Dependency Associated with LLM Chatbots

Author(s)	Year	Study Focus	Key Findings
Kreijkes et al.	2025	Experimental comparison of LLM chatbot support versus traditional note-taking for retention	Found that while students reported a positive experience with the chatbot, it was less effective for information retention compared to traditional note-taking methods.
Bastani et al.	2024	Experimental study into effects of different LLM chatbot support versus no support	LLM support improved performance during practice, but performance declined below the level of students who never used the chatbot once it was removed. This effect was mitigated when the LLM was trained with prompts to act as a tutor designed to safeguard learning.
Darvishi et al.	2024	Experimental study into impact of LLM chatbot support on the quality of peer feedback and student agency	Found that chatbot support enhanced performance and quality of feedback, which declined when it was removed, suggesting student dependence and "distributed metacognition" (p.10).

Note on acronym: LLM = Large Language Model.

chatbots in their original form. As AI chatbots can have an infinite combination of roles, training, datasets, and outputs, their impacts also vary considerably. It may be that only certain well-designed and effectively used LLM chatbots or chatbot tutors may truly enhance student experiences and outcomes. However, what constitutes “well-designed” and “effectively used” in this context remains ambiguous.

In addition, the sophisticated neural networks and stochastic nature of LLM chatbots complicate the assessment of their impact, as they unpredictably generate varied responses to the same question asked by different users (Tlili et al., 2023). This, combined with the pace of development in the field, has turned LLM chatbots into moving targets which independent academic research is ill-equipped to hit.

Moreover, as highlighted by Lehmann, Cornelius and Sting (2024) and implied by Bastani et al. (2024), learning with LLM chatbots relies on productive use and cognitive engagement. Thus, one explanation for the discrepancies in findings may stem from the fact that the amount a student learns likely depends on their level of cognitive engagement, which will depend on the student, the chatbot, and the specific use, in turn influenced by external factors, such as time pressure and other demands (Abbas, Jam, & Khan, 2024). In their study, Lehmann et al. (2024) found that positive learning outcomes were reserved for students who used LLM chatbots as tutors, rather than those who relied on

them to solve problems for them with minimal effort. They also noted that students with less experience using chatbots tended to use these tools less productively.

These findings suggest that teaching students effective ways to use LLM chatbots and providing opportunities to gain experience could enhance productive uses and cognitive engagement. However, positive learning outcomes also depend on chatbots giving students accurate information, which LLM chatbots do not guarantee. For instance, in Sun et al. (2024), while improved learning was observed in tasks supported by chatbots, the overall performance of students using the chatbot was no better than that of those without it, primarily due to inaccurate information generated by the chatbot. Hallucinations may therefore be a significant barrier to learning performance using LLMs.

For an overview of these studies, see Table 3.

3.5 Beyond performance: What more could LLM chatbots offer?

Many studies discussed here, examining the impact of LLM chatbots on learning, have focused on short-term effects and surface-level learning, such as retention and test performance. This suggests that negative findings regarding AI's impacts on learning may primarily indicate a detrimental effect on surface-level, rather than deep, learning.

Table 3. Studies Identifying Factors Contributing towards Inconsistent Outcomes in LLM Chatbot Research

Author(s)	Year	Study Focus	Key Findings
Phung et al.	2023	Experimental study into impact of different ChatGPT versions	Found ChatGPT-4 significantly outperformed ChatGPT-3.5 in computer programming support, nearly matching the impact of a human tutor.
Tlili et al.	2023	Three-part study investigating user experience	Found significant variation in ChatGPT responses to the same question for different users.
Lehmann, Cornelius & Sting	2024	Experimental study into learning with LLM chatbots	Found that learning outcomes were more positive for students who used LLM chatbots as tutors and engaged cognitively.
Abbas, Jam, & Khan	2024	Correlative study of student self-report questionnaires	Noted correlative relationships between higher time and workload pressure, ChatGPT use, and higher procrastination, memory loss, and lower grades.
Sun et al.	2024	Experimental study into LLM chatbots	Observed improved learning in tasks supported by ChatGPT, while overall performance was not better than control groups due to inaccurate information provided by the chatbot.

Note on acronym: LLM = Large Language Model.

The distinction between surface learning - where students retain facts or procedures without deeper understanding - and deep learning, which involves cognitively engaging with content and connecting it to prior knowledge, stems from foundational work by Marton and Säljö (1976). While surface learning has merit and is valuable for facilitating deep learning, the latter engages complex cognitive functions such as analysis, evaluation, and synthesis, resulting in a significant transformation of knowledge. This process enables students to pursue personal growth and actively engage with learning materials, linking new information to their prior experiences (Desierto, De Maio, O'Rourke, & Sharp, 2018). Deep learning is often associated with "higher quality learning outcomes" (Marton & Säljö, 1976, p. 4); however, it is more challenging to measure, encompassing behaviours, skills, and values across disciplines, suggesting that a high test score may reflect surface learning rather than true understanding (Marton & Säljö, 1976).

This distinction is significant: by providing instant access to knowledge, LLM chatbots may reduce the need to retain specific information, leading to less effective surface learning and retention - a phenomenon consistent with the implications of other technology known as the 'Google effect' (Sparrow, Liu, & Wegner, 2011). However, this does not negate the potential of these tools to support deep learning. From a social learning theory perspective, for example, chatbots may serve as a knowledgeable presence essential for deep learning. Bandura (1977) asserts that learning occurs through the observation and modelling of actions and thought processes, enabling learners to internalise strategies. Similarly, according to Vygotsky (1978), learning is fundamentally a social process, where dialogue and interaction with more knowledgeable others, such as teachers and peers, is essential for the development of higher-order thinking skills, particularly within the Zone of Proximal Development (ZPD), where these individuals facilitate learners' understandings of complex concepts. As LLM chatbots excel in providing knowledge from extensive datasets and possess social and conversational features, they are capable of modelling actions, aiding understanding, and offering scaffolding for learners. Additionally, as LLM chatbots have gained in sophisticated multimodal abilities, voice engagement arguably enables them to become more 'conversational' and lowers barriers to use. Consequently, LLM chatbots may be well-equipped to facilitate deep learning.

Building on this idea, Mercer, Wegerif, and Major (2019) and Wegerif (2019) argue that dialogue in particular is a key mechanism through which learners co-construct knowledge and understanding and deeper learning is achieved. Therefore, as Wegerif and Major (2024) highlight, by

creating a dialogic space, LLM chatbots have the potential to deepen students' learning, enabling them to enhance their understanding, reflect, ask questions, engage in discussions, and encounter new perspectives, insights, and challenges. However, realising these benefits is not automatic and would likely rely on either well-designed chatbots or well-prepared students able to use LLM chatbots effectively as a dialogic tool. Indeed, as it is well established that students must learn how to benefit from dialogue - even among humans (Alexander, 2008) - it follows therefore that teaching students how to engage in dialogue with LLM chatbots would be equally crucial to unlocking their full potential.

In these ways, LLM chatbots may open new possibilities for education and deep learning that extend beyond surface-level achievements. As Wegerif and Major argue, from a philosophy of education perspective, there is an essential distinction between 'training' and 'education' (2024, p.137). They argue that much of what is labelled 'education', which prioritises rote tasks - such as examinations and essays - amounts to 'training', or conditioning students to produce responses which LLM chatbots can easily replicate. Training, therefore, is closely aligned with surface-level learning, whereas genuine education involves going further, learning to think critically and can be considered deep learning. Wegerif and Major contend that technology, including LLM chatbots, can help cultivate deeper learning through dialogue and dialogic spaces. Therefore, while LLM chatbots may adversely affect surface-level 'training', they may also hold significant potential to enhance education.

The impact of LLM chatbots in education is therefore complex and nuanced, prompting further exploration into what education should offer students and how it should be measured. As Wegerif and Major (2024) note, AI tools now perform tasks traditionally associated with assessing student learning - such as writing essays and rote memorisation - challenging the necessity of these skills and the relevance of conventional assessment methods. This reality necessitates a re-evaluation of the educational system alongside the integration of technology, prompting an important examination of the role of LLM chatbots in the learning experience.

4. Mastery learning behaviours in education

In exploring how LLM chatbots might contribute to more meaningful learning experiences like deeper learning, an important yet largely overlooked consideration in these studies into the impact of LLM chatbots is the learning behaviours associated with more positive outcomes. While learning behaviours are vast, this paper focuses on four fundamental

learning behaviours, also referred to as ‘mastery’ behaviours: perseverance, challenge-seeking, effort, and resilience. According to Porter et al. (2020), ‘mastery behaviours’ are distinct yet connected behaviours. Together the term refers to “differences in how children respond to challenge, effortful work, and failure” (p.6.). Individually, they can be defined as follows:

- *Perseverance*: According to the Cambridge Dictionary, perseverance is “continued effort to do or achieve something, even when this is difficult or takes a long time” (Cambridge Dictionary. n.d.). Perseverance can be understood as persisting when faced with a challenging activity (Kleitman et al., 2022) or “sustained engagement with a difficult task” (Porter et al., 2020, p.6.) and is therefore often measured in time spent working at a difficult activity.
- *Challenge-seeking*: Challenge-seeking can be defined as seeking and embracing challenges in learning. In education research it is sometimes operationalised as “selecting a more difficult task instead of an easier one” (Porter et al., 2020, p.6).
- *Effort*: Effort can be defined as conscious exertion of mental energy towards achieving a goal or completing a task. Porter et al. (2020) define effort simply as “investment in learning” (p.6), while the Collins Dictionary refers to notions of “trying hard” (Collins English Dictionary, n.d). Objective effort is sometimes operationalised as “using more mental energy for any given unit of time” (Dunlosky, Badali, Rivers, & Rawson, 2020, p.1164). Effort in this context is quantitative and qualitative, it involves not only working for longer but also engaging greater cognitive resources (i.e., by producing more output and of higher quality).
- *Resilience*: Resilience can be defined as “the ability to be happy, successful, etc. again after something difficult or bad has happened” (Cambridge Dictionary, n.d.-b) or the “capacity to recover after failure” (Porter et al., 2020, p.6). In education research, it is sometimes operationalised as students’ capacity to recover quickly or ‘bounce back’ to pre-challenge performance levels after a failure or a very difficult task.

According to research and literature, these four mastery behaviours have an important and positive impact both within education and in life more broadly. Within education, these mastery behaviours are associated with greater academic success (Bittmann, 2021; Dweck, 2006; Duckworth, Peterson, Matthews, & Kelly, 2007; Kaya & Karakoc, 2022; Porter et al. 2020; Sulik, Finch & Obradović, 2020; Yin, Heard, Szabo, & Ankoor, 2021), more productive problem

solving and learning strategies (Burnette et al. 2013; Meyer, Turner & Spencer, 1997; Porter et al., 2020), increased learning, competence, and perceived competence (Meyer, Turner & Spencer, 1997), more positive attitudes toward schooling and engagement (Xenofontos & Mouroutsou, 2023), and even lower dropout intentions (Bittmann, 2021).

Beyond education, research suggests greater resilience (Bittmann, 2021; Xenofontos & Mouroutsou, 2023), grit (“perseverance and passion for long-term goals” (Duckworth et al., 2007, p.1087) (Datsu, 2021), and challenge-seeking (Abuhamdeh & Csikszentmihalyi, 2012; Dweck, 1986; 2006; Mouratidis, Michou, Demircioğlu, & Sayil, 2018) are associated with a range of positive life outcomes, such as improved physical and mental health, wellbeing, self-efficacy, enjoyment, fulfilment, life satisfaction, overall happiness, positive habits, and career outcomes. These behaviours are therefore important for shaping traditional educational outcomes, such as grades, but also as educational outcomes in their own right.

While it is important to recognise that mastery behaviours are just one of many factors contributing to positive outcomes - and to avoid overstating their impact - research does indicate that they positively impact individuals both in their education and in their wider lives. Therefore, students would significantly benefit from an education system that fosters and supports the productive development of these behaviours. Increasing productive behaviours is thus a vital component of an education system that empowers students for success.

Additionally, these behaviours may be more crucial than ever in preparing students for the future. As AI advances, it is expected to outperform humans in many areas, potentially displacing traditional core educational skills (Tuomi, 2019). In this context, ‘future’ or 21st-century skills will gain prominence over mere knowledge (Mogavi et al., 2024). Soft skills, as emphasised by the OECD (2019), will evolve into core skills, with qualities like perseverance and resilience being essential for individuals to adapt to changing environments. Therefore, while LLMs may challenge the role of rote learning, examinations, and written essays, and raise fundamental questions about what education should teach and how, they may ultimately reinforce, rather than diminish, the importance of an educational system that fosters productive mastery learning behaviours. A crucial question. Therefore, is how education can nurture these behaviours and whether LLM chatbots have the potential to cultivate such valuable qualities in students.

In exploring this potential, Social Cognitive Theory (SCT), Behaviourist Theory, and Self-Determination Theory (SDT) provide essential insights into behaviour development that help to understand how education can influence positive behaviours. SCT posits that behaviours are learned through observation, modelling, and the social reinforcement of attitudes and action, and are further shaped by self-efficacy, as individuals' beliefs in their capabilities influence the actions they undertake (Bandura, 1977). Behaviourist Theory highlights the role of external stimuli and feedback in reinforcing behaviours, illustrating how positive experiences can encourage positive behavioural cycles (Skinner, 1953). Meanwhile, SDT emphasises the importance of fulfilling the intrinsic needs of autonomy, competence, and relatedness, which are crucial for sustaining effort and engagement (Deci & Ryan, 2000). Therefore, by addressing important factors highlighted by SCT, Behaviourism, and SDT, learning environments may significantly influence student behaviour - a theoretical foundation that sets the stage for exploring the potential impact of LLM chatbots on these key behaviours in Section 5.

5. Could LLM chatbots increase mastery behaviours?

As highlighted in Section 3, findings regarding the impact of LLM chatbots on learning and performance are mixed. This variability is likely because the impact on learning will significantly depend on factors such as the specific chatbot used, the subject matter, its pedagogical capabilities, content knowledge, and how effectively it adapts to individual students. Consequently, these findings may quickly become outdated as chatbots evolve. This paper therefore shifts the focus to consider thoughtfully the impact of LLM chatbots on student behaviours as important educational outcomes in their own right, proposing that key features of LLM chatbots - such as their consistent availability, instant feedback, dialogic and relational qualities, and student-directed interactions - may positively influence key learning behaviours. Examining this relationship is important not only because these behaviours contribute towards success and future preparedness, but also therefore because the behavioural effects may be more likely to derive from key features of LLM chatbots, making them less sensitive to variations among individual chatbots than traditional learning outcomes might be. Moreover, these perspectives prompt us to thoughtfully consider what individuals can achieve in partnership with AI, expanding the focus beyond what individuals can accomplish on their own. This progression in understanding recognises that human+AI collaborations and their unique outcomes also offer valuable contributions

to education - an outlook that may be necessary if education and assessment are to be authentically aligned with the real world and meaningful in preparing students for the future (Lodge et al., 2023; Pratschke, 2024; Watson, 2025).

Despite growing research on LLM chatbots in education, however, the relationship between LLM chatbots and student learning behaviours, in particular mastery behaviours, remains largely unexplored. In Xu, Wang, Zhang, and Zheng (2024), interviews with higher education experts in China indicated their perceptions that ChatGPT could foster non-cognitive skills like motivation, self-efficacy, and self-determination, as well as behaviours such as perseverance and self-regulation, highlighting the psychological and behavioural implications discussed in this section. However, a focused database search revealed no existing research specifically examining the relationship between chatbots (LLM or otherwise) and key mastery learning behaviours, such as resilience, effort, perseverance, or challenge-seeking, highlighting a significant gap in understanding.

This paper argues that LLM chatbots could foster more productive mastery behaviours primarily through the key features underlying their support. Although chatbots do not replace other forms of assistance, the support they offer presents distinct advantages over finite teacher support, textbooks, and search engines. For instance, LLM chatbots are unique in providing consistently available, 1:1 tailored responses and personalised feedback (Gupta & Chen, 2022; Yildiz Durak, 2022). Moreover, LLM chatbots offer more than just answers; they engage in dialogue (Mogavi et al., 2024). This dialogic quality not only has the potential to enhance learning experiences (Wegerif & Major, 2024), but also to create a relational environment where some students feel they are collaborating with a knowledgeable study partner attuned to their needs (Antony & Ramnath, 2023; Hmoud et al., 2024). By providing practical support that helps students persist through challenges and develop confidence while also creating a positive learning environment that promotes motivation, autonomy, and a sense of safety, key features of LLM chatbots - such as the consistent availability, instant feedback, and student-directed interactions - could foster productive mastery learning behaviours.

5.1 Psychological and behavioural effects of LLM Chatbots' practical learning support

Firstly, based on Social Cognitive Theory and Behaviourist Theory, the practical support provided by LLM chatbots may influence more productive mastery learning behaviours. Research shows that individuals frequently turn to LLM chatbots to tackle difficult tasks or concepts (Impact Re-

search, 2024; Mogavi et al., 2024). By modelling strategies and offering tailored and immediate support during difficulties, LLM chatbots may thereby enable students to persist in tasks rather than giving up when they encounter challenges. This ongoing support could make these challenges feel less overwhelming, more manageable, and within the individual's capabilities. By virtue of knowing the chatbot can assist with these challenges, the practical support may lower the perceived difficulty of tasks, reducing associated negative behaviours, like quitting prematurely or avoiding challenges altogether (Baten et al., 2020). Indeed, in a study by Kumar, Rothschild, Goldstein, and Hofman (2023) involving 1,200 participants solving mathematical problems under different conditions, some given direct answers, while others received LLM-generated explanations, those who received explanations from the chatbot reported finding the problems less difficult than those who were shown the answers directly. This suggests that LLM-generated explanations, by clarifying difficult concepts, can reduce the perception of a task's difficulty and encourage students to persevere through challenges.

Moreover, by enabling students to complete more challenging tasks, LLM chatbots may help establish a positive learning cycle that encourages productive mastery behaviours by creating a more rewarding experience. For example, by allowing students to experience the rewards of persisting, behaviourist principles would suggest chatbots may make students more likely to continue engaging in positive behaviours (Yang, Chen, Wang, & Chen, 2023).

Furthermore, in allowing students to complete increasingly challenging tasks, LLM chatbots may also enhance confidence and self-efficacy, which could lead to more productive behaviours, given that higher self-efficacy is linked to mastery behaviours like effort and perseverance when facing challenges (Hayat, Shateri, Amini, & Shokrpour, 2020), persistence (Li-Fang, 2006; Tinto, 2022), and resilience (Cassidy, 2015). The argument that LLMs can enhance confidence and self-efficacy aligns with various studies. For example, in qualitative research by Hasanein and Sobaih (2023) and Dai et al. (2023), participants in diverse educational settings in Saudi Arabia and Australia reporting thinking ChatGPT could help enhance confidence. In experimental studies, Kumar et al. (2023) found that the participants who received LLM-generated explanations for mathematical questions reported greater confidence in their learning and performance, while Yilmaz and Yilmaz (2023) found that ChatGPT support during coding practice significantly increased self-efficacy among Turkish undergraduates. Similarly, Chang, Huang, and Gau (2022) and Lee, Hwang, and Chen (2022) found that pre-ChatGPT chatbots boosted

Taiwanese students' self-efficacy compared to conventional teaching methods when used in nursing training or for after-class reviews - an effect both studies attributed to chatbot feedback and dialogic responses.

In AI research connecting these improvements to behavioural outcomes, Wu et al. (2023) reported increased self-regulated learning among Taiwanese students using ChatGPT, which they claimed was mediated by enhanced self-efficacy. However, this remains the only study linking these traits directly with behaviour, and none have yet explored their relationship to mastery behaviours. While pre-ChatGPT, Yildiz Durak (2022) found that a chatbot used in a visual design course did not improve engagement for Turkish students, and Fryer et al. (2020) observed no improvement in self-efficacy with chatbot partners among Japanese university students, LLM chatbots may generally be more effective in this regard due to the more consistent characteristic high-quality of output they produce, which distinguishes them from all prior chatbots.

5.2 Psychological and behavioural effects of LLM Chatbots' positive learning environment

Secondly, drawing on Self-Determination Theory, chatbots could increase productive behaviours by creating a positive learning environment which is simultaneously autonomy-supportive, motivational, and safe.

5.2.1 An autonomy-supportive tool

Within educational literature, autonomy, self-determination, and enhanced motivation and engagement are interrelated factors linked to more productive behaviours (Deci & Ryan, 2000), such as challenge-seeking (Baten et al., 2020; Harter, 1981; Mierke, Scheidtmann & Ibrahimova, 2017), perseverance, and effort (Tinto, 2022). In particular, research suggests that autonomy-supportive teaching practices positively impact student motivation, engagement, and resilience (Ma, 2021). Therefore, if LLM chatbots offer an autonomy-supportive tool that increases motivation and engagement, they may well promote greater mastery behaviours. Although the relationship between these three factors remains empirically underexplored, two key reasons emerge as to why LLM chatbots could enhance motivation and autonomy: first, by allowing students to independently direct their interactions with the chatbot, and second, by providing a learning experience that is both relational and interactive.

Firstly, LLM chatbots could enhance autonomy and motivation as they are directed by the student. This allows students to pursue individual interests and access diverse and engaging presentations of information. Additionally, this enables students to independently seek support and feedback, providing greater autonomy and control which could enhance self-determination and ownership of learning (Dai et al., 2023); these are factors which, according to self-determination theory, are also linked to greater motivation (Deci & Ryan, 2000). Indeed, several studies suggest LLM chatbots can increase task enjoyment (Hmoud et al., 2024), engagement (Antony & Ramnath, 2023; Khurma et al., 2024), and motivation (Mogavi et al., 2024; Yilmaz & Yilmaz, 2023); motivational effects which may serve as vehicles through which LLM chatbots promote positive learning behaviours.

However, the impact of LLM chatbots on self-determination remains debated. By assuming responsibility for outcomes and reducing the need to engage cognitively with work, Abbas, Jam and Khan (2024) suggest that LLM chatbots might reduce student ownership and feelings of accomplishment, Mogavi et al. (2024) question whether LLM chatbots reduce students' autonomy and competence, fostering more passive learning behaviours, and Luo et al. (2024) highlight concerns among practitioners that easy access to answers could weaken self-determination, interest, and motivation. Arguably, as previously noted, positive outcomes may depend on using ChatGPT to support learning and growth rather than as a shortcut; if students merely copy responses, self-determination and motivation could diminish. Following this argument, LLM chatbots designed to play the role of tutor and guide students without providing answers may be better positioned to foster positive behaviours. However, no research has yet empirically tested this association.

A further means by which LLM chatbots may increase motivation and engagement, thereby enhancing positive behaviours, is through the relational experience they can provide, where students may feel that they are being helped by or collaborating with a partner, friend, or teammate, making the experience more interactive and fun. While caution may be necessary in presenting LLM chatbots as anything more than 'tools', as, by referring to these tools as 'partners', it is possible to incorrectly extend the trust and goodwill typically reserved for genuine friends or allies (Deneen, 2024). However, independent of terminology, research shows that students often do develop a sense of partnership and affinity with LLM chatbots (Antony & Ramnath, 2023). Therefore,

while this label may not be desirable, if LLM chatbots are perceived and used as a 'presence' more than merely a tool, it is this psychological response and perspective which should be considered when exploring their impact (Pratschke, 2024).

The argument that LLM chatbots create more positive experiences is supported by research. For example, while Kreijkes et al. (2025) found LLM chatbot support had a less positive impact on subsequent retention performance, they also found that students self-report perceptions of the experience overwhelmingly indicated more positive feelings towards using the LLM chatbot with their learning than learning alone. Consequently, LLM chatbots may create a more positive learning experience that could result in more productive mastery behaviours. While these chatbots may have no positive effect on surface learning in the short term, these positive experiences could potentially contribute to more productive learning behaviours and even deeper learning over time; however, for this it is essential to consider whether the positive experience would endure. While studies such as Hmoud et al. (2024) found ChatGPT had a sustained impact in student interviews over a 3-6-month period, Fryer et al.'s (2017) earlier research on a pre-ChatGPT chatbot noted the motivational novelty effect of new technology in education, which subsides over time, meaning any motivational benefits and subsequent behavioural outcomes could be short-lived; this is a consideration which should be explored.

5.2.2 Safe learning environment

Lastly, LLM chatbots could increase productive behaviours by creating a safe learning environment. As highlighted by Abdillah, Partino, and Madjid (2023), chatbots could help alleviate student anxiety. Indeed, LLM chatbots' immediate, tailored assistance could help reduce feelings of helplessness, incompetence, frustration, and overwhelm in complex problem-solving and preventing rumination by promptly correcting errors (Hawanti & Zubaydullovna, 2023). For many, the relational experience they provide may enhance feelings of safety and comfort, yet, as an added bonus, as artificially intelligent entities, they also enable students to bypass fears of appearing unintelligent or burdensome by making mistakes or asking 'silly' questions (Gupta & Chen, 2022; Haque & Rubya, 2023; Terblanche, Molyn, Williams, & Maritz, 2023).

By serving as a safety net and allowing students to engage more productively with learning, LLM chatbots could reassure and empower students, allowing them to step out of their comfort zones and engage in more productive learning behaviours, such as prolonged persistence and challenge-seeking. Indeed, in Lee, Hwang and Chen (2022), authors argued that Taiwanese students using chatbots for after-class reviews asked more questions and engaged in deeper conversations which authors argued was likely due to feeling less judged, indicating the importance of this consideration when exploring how LLM chatbots might influence student behaviours.

6. How well equipped are LLM Chatbots to promote behaviours?

Despite potential, universal limitations, such as their non-human nature, baked-in bias and their ability to reduce effort required to produce work, challenge the potential of LLM chatbots to influence productive learning behaviours.

6.1 The non-human factor

LLM chatbots are inherently non-human. While this offers certain benefits, it also presents challenges. For example, due to the absence of embodied and affective experience (Sharples, 2023), LLM chatbots may lack human empathy and understanding (Antony & Ramnath, 2023; Khurma et al., 2024). ChatGPT responses and feedback can be rigid and inflexible (Gupta & Chen, 2022; Kuhail, Alturki, Alramlawi, & Alhejori, 2023), “dull, monotonous, and cold” (Luo et al., 2024, p.107) and inferior to those from human-teachers (Gattupalli, Malloy, & Edwards, 2023).

As Polyportis (2024) highlights, affective dimensions of technology use are an important, yet often overlooked, aspect, and negative interactions are likely to have a detrimental impact. Students for whom the experience of LLM chatbots is negative are unlikely to benefit from increased engagement, motivation, self-efficacy, productive use and outcomes, behavioural or other.

However, the hindrance posed by their artificially intelligent nature might not be as significant as suggested. Escalante, Pack and Barrett (2023) conducted a six-week experiment comparing written feedback from AI and human tutors and found no significant difference in students’ learning outcomes and an almost equal preference for both. As Sun et al. (2024) found, students’ perceptions of these tools can become more positive over time. Given that reports indicate that the use of tools like ChatGPT increased

between 2023 and 2024 (Impact Research, 2024) and recent surveys showing a majority of students reporting a positive view of LLM chatbots (Das & Madhusudan, 2024; Impact Research, 2024; National 4-H Council & Hart Research, 2024), as models improve student experiences may be becoming increasingly positive over time.

Indeed, given the emphasis on developing more sophisticated and human-like communication abilities, including verbal interaction and emotion recognition, LLM tools like ChatGPT are rapidly becoming more advanced and equipped to engage in social roles. This is evidenced in the research. For example, while in 2023 Elyoseph et al. reported that ChatGPT-3.5 demonstrated higher emotional awareness than the general population, in 2024 Sufyan, Fadhel, Alkhathami, and Mukhadi found that ChatGPT-4 scored notably higher than human psychologists on the standard Social Intelligence Scale. A recent Stanford University study (Mei, Xie, Yuan, & Jackson, 2024) revealed that ChatGPT-4 exhibited behavioural tendencies indistinguishable from humans, surpassing ChatGPT-3.5 in human-likeness and deviating only from average human behaviour to behave with greater altruism and cooperation, prioritising maximum payoff for both itself *and* its human partner. And in a study published in May 2024, Strachan et al. (2024) found ChatGPT-4 performed equivalent or superior to humans in theory of mind tasks. This is not to say that having AI chatbots to fulfil all human roles is desirable – this is a separate debate – but to say that, when considering the experience of students, the ability of these artificially intelligent tools to bring about positive affect and behaviours may be improving.

6.2 Bias and inequities in LLM outputs

While LLMs may be well-equipped to adapt to students and contexts (Chiu, Moorhouse, Chai, & Ismailov, 2023; Khurma et al., 2024; Tlili et al., 2023; Zamfiroiu et al., 2023) potentially mitigating human-teacher bias (Hsu, Huang, Hwang, & Chen, 2023), due to their design, training data, and neural network processing, they remain highly biased (Kasneci et al., 2023; Kohnke, Moorhouse, & Zou, 2023; Zhu et al., 2023). Such baked-in biases affect decisions regarding language, learning materials, and support, potentially favouring those most closely aligned with the technology’s design, data, and training, thereby creating unequal educational experiences and opportunities (Baker & Hawn, 2021; Karumbaiah, Ocumpaugh, & Baker, 2021). Indeed, a recent study by the Stanford Institute of Human-Centered Intelligence (Shieh, Vassel, Sugimoto, & Monroe-White, 2024; Vassel, Shieh, Sugimoto, & Monroe-White, 2024) highlights the significant magnitude of bias in LLM outputs, particularly in terms of erasure – the absence

of certain stories and perspectives, and subordination - where marginalised identities are portrayed as inferior, and negative stereotyping. The study, which analysed 150,000 LLM-generated outputs, found that, statistically, individuals from minoritised and marginalised groups are hundreds to thousands of times more likely to encounter outputs that subordinate their identities rather than represent or empower them. These outputs often perpetuate harmful stereotypes which are known to cause psychological harm, such as the low attaining Latina student, 'Maria', or high achieving Asian STEM student, 'Priya'. The authors highlight the lack of research on this issue and stress the need for a critical investigation of the psychosocial effects of the biases present in the text outputs of leading generative AI models.

As Vassel et al. (2024) argue, if LLM chatbots cannot universally ensure positive experiences, affect, or outcomes, they are unlikely to universally promote positive behaviours. Ironically, those who could benefit most from educational chatbot support may be those least likely to be represented in the data and training of LLMs, potentially further marginalising their experience. While studies do indicate that ChatGPT can adapt effectively to various roles, contexts, and partners when prompted (Gattupalli, Malloy, & Edwards, 2023; Mei et al., 2024), suggesting that carefully trained tutors could support diverse students if mindfully trained to do so, as empirical research on bias in LLM tutors, how it can be corrected, and the direct implications for students is sparse, this remains unclear.

6.3 'Lazy' learning

Finally, as Abbas, Jam and Khan (2024) and Kiikerpill and Siibak (2023) highlight, as ChatGPT can produce work with minimal human effort, concern surrounds its potentially negative impact on student learning and behaviours, including decreased cognitive engagement and effort, greater indolence, and more superficial learning habits (see also Mogavi et al., 2024). Thus, LLM chatbots might hinder rather than enhance behavioural productivity.

In their study, Abbas, Jam and Khan used a time-lagged design to survey 494 university students in Pakistan over three time points (T1, T2, and T3) spaced 1 to 2 weeks apart. T1 represented the initial measurement, where students' self-reported time and workload pressure were assessed. T2 referred to the second time point, capturing the extent of self-reported ChatGPT use following the pressures reported at T1. T3, the final stage, examined

self-reported outcomes, such as procrastination, memory loss, and academic performance. Their findings showed that students experiencing higher time and workload pressure at T1 reported increased ChatGPT use at T2, which correlated with higher procrastination, memory loss, and lower grades at T3. However, as the study lacked experimental design and factors at T2 and T3 could have existed previously, this does not show ChatGPT *caused* negative consequences; perhaps students facing greater pressure and academic challenges were simply more inclined to seek ChatGPT assistance.

In contrast, as noted above, scholars such as Wegerif and Major (2024) contend that tools like ChatGPT could enhance deeper engagement with learning and the development of skills (Mogavi et al., 2024), a claim that is empirically supported by various studies. For instance, Yilmaz and Yilmaz (2023) found that using ChatGPT for coding increased Turkish undergraduate students' creative and computational thinking as well as problem-solving behaviours. Australian postgraduate students in Dai et al. (2023) reported that ChatGPT improved critical thinking. Vasconcelos and dos Santos (2023) conducted a case study in the Brazilian context showing that using ChatGPT and Bing Chat as thinking tools supported reflective, creative, and critical thinking, and problem-solving skills in STEM subjects. Additionally, Essel et al. (2024) found that integrating ChatGPT into a Ghanaian undergraduate course significantly increased students' critical, reflective, and creative thinking skills compared to a control group.

Moreover, studies, such as Lee et al. (2024), Bastani et al. (2024), and Kumar et al. (2023), show that enhancing ChatGPT with additional features which encourage students to try independent problem-solving and offer guidance instead of answers ensures more productive cognitive engagement and behaviour than ChatGPT alone. As LLM tutors can be programmed specifically to facilitate students' cognitive engagement, they can remove the quick shortcut aspect provided by ChatGPT (Khan, 2024), potentially making them more likely to foster productive behaviours.

In summary, there are important limitations which not only call into question the use of LLM chatbots in education, but could also challenge the capacity of LLM chatbots to influence positive learning behaviours. However, questions remain about whether proposed limitations are indeed limitations, which models and students they affect, whether they can be overcome, and whether they ultimately hinder the impact of LLM chatbots on student learning behaviours.

7. Conclusion

Despite the importance of understanding both benefits and drawbacks for effective integration (Miao et al., 2021), given the field infancy and pace of academic research, the picture painted on the impact of LLM chatbots in education continues to present gaps, with extremely limited studies into their relationship with student learning behaviours.

Drawing on the latest research, this literature review highlights that LLM chatbots have potential in education as well as presenting significant challenges. It advocates for a critical approach to integration, highlighting the importance of empirically exploring all of the potential risks and benefits these tools bring. It suggests that an important yet unexplored impact of LLMs may be the impact on important learning behaviours, like effort, challenge-seeking, perseverance, and resilience. However, while emerging research indicates a range of means by which LLM chatbots may lead to more positive behaviours, these tools also raise concerns and may be limited in their capacity to foster these productive behaviours. Ultimately, very little has been done to empirically explore changes to students' real-time learning behaviours, meaning this impact is yet to be understood. Given the discussions and research presented in this review, this paper highlights and concludes that investigating the impact of LLM chatbots on learning behaviours holds significant potential and demands urgent attention in the evolving educational landscape.

References

- Abbas, M., Jam, F.A. & Khan, T.I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, 21, 10. <https://doi.org/10.1186/s41239-024-00444-7>
- Abdillah, H. Z., Partino, P., & Madjid, A. (2023). Enhancing Student Well-being through AI Chat GPT in the Smart Education University Learning Environment: A Preliminary Review of Research Literature. In: *E3S Web of Conferences*, 440, 05005. <https://doi.org/10.1051/e3sconf/202344005005>
- Abuhamdeh, S., & Csikszentmihalyi, M. (2012). The importance of challenge for the enjoyment of intrinsically motivated, goal-directed activities. *Personality and Social Psychology Bulletin*, 38(3), 317-330. <https://doi.org/10.1177/0146167211427147>
- Alexander, R. J. (2008). *Towards Dialogic Teaching: rethinking classroom talk (4th Edition)*. Dialogos.
- Alneyadi, S., & Wardat, Y. (2023). ChatGPT: Revolutionizing student achievement in the electronic magnetism unit for eleventh-grade students in Emirates schools. *Contemporary Educational Technology*, 15(4), ep448. <https://doi.org/10.30935/cedtech/13417>
- Antony, S., & Ramnath, R. (2023). A Phenomenological Exploration of Students' Perceptions of AI Chatbots in Higher Education. *IAFOR Journal of Education*, 11(2), 7-38. <https://doi.org/10.22492/ije.11.2.01>
- Baker, R. S., & Hawn, A. (2021). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Bandura, A. (1977). *Social learning theory*. Prentice-Hall.
- Bastani, H., Bastani, O., Sungu, Alp., Ge, H., & Kabakci, Ö., & Mariman, R. (2024) Generative AI Can Harm Learning. *The Wharton School Research Paper*. Accessible at: <https://ssrn.com/abstract=4895486>. <https://doi.org/10.2139/ssrn.4895486>
- Baten, E., Vansteenkiste, M., De Muynck, G. J., De Poortere, E., & Desoete, A. (2020). How can the blow of math difficulty on elementary school children's motivational, cognitive, and affective experiences be dampened? The critical role of autonomy-supportive instructions. *Journal of Educational Psychology*, 112(8), 1490–1505. <https://doi.org/10.1037/edu0000444>
- Bittmann, (2021). When problems just bounce back: about the relation between resilience and academic success in German tertiary education. *SN Social Sciences*, 1(2), 65. <https://doi.org/10.1007/s43545-021-00060-6>
- Burnette, J. L., O'boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139(3), 655. <https://doi.org/10.1037/a0029531>
- Cambridge Dictionary. (n.d.-a). Perseverance. *Cambridge University Press*. Accessible at : <https://dictionary.cambridge.org/dictionary/english/perseverance?q=perseverance>

- Cambridge Dictionary. (n.d.-b). Resilience. *Cambridge University Press*. Accessible at: <https://dictionary.cambridge.org/dictionary/english/resilience>
- Cassidy, S. (2015). Resilience building in students: The role of academic self-efficacy. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01781>
- Chang, C. Y., Hwang, G. J., & Gau, M. L. (2022). Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *British Journal of Educational Technology*, 53(1), 171-188. <https://doi.org/10.1111/bjet.13158>
- Chiu, T. K., 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*, 1-17. <https://doi.org/10.1080/10494820.2023.2172044>
- Collins English Dictionary. (n.d.). Effort. *Collins*. Accessible at: <https://www.collinsdictionary.com/dictionary/english/effort>
- Dahlkemper, M. N., Lahme, S. Z., & Klein, P. (2023). How do physics students evaluate artificial intelligence responses on comprehension questions? A study on the perceived scientific accuracy and linguistic quality of ChatGPT. *Physical Review Physics Education Research*, 19(1). <https://doi.org/10.1103/physrevphyseducres.19.010142>
- Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology*, 39(4), 74-88. <https://doi.org/10.14742/ajet.8843>
- Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967. <https://doi.org/10.1016/j.compedu.2023.104967>
- Das, S. R., & Madhusudan, J. V. (2024). Perceptions of Higher Education Students towards ChatGPT Usage. *International Journal of Technology in Education*, 7(1), 86-106. <https://doi.org/10.46328/ijte.583>
- Deci, E. L. & Ryan, R. M. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1), 54-67. <https://doi.org/10.1006/ceps.1999.1020>
- Deneen, C. (2024). AI and assessment in higher education: Problems, possibilities, and pathways. *Pacific Journal of Technology Enhanced Learning*, 6(1), 28. <https://doi.org/10.24135/pjtel.v6i1.199>
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2024). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Desierto, A., De Maio, C., O'Rourke, J., & Sharp, S. (2018). Deep or Surface? The learning approaches of enabling students in an Australian public university. Paper presented at the STARS (Students Transitions Achievement Retention and Success) Conference, Auckland, New Zealand. Accessible at: https://www.researchgate.net/publication/353837764_Deep_or_Surface_The_learning_approaches_of_enabling_students_in_an_Australian_public_university
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087-1101. <https://doi.org/10.1037/0022-3514.92.6.1087>
- Dunlosky, J., Badali, S., Rivers, M. L., & Rawson, K. A. (2020). The role of effort in understanding educational achievement: Objective effort as an explanatory construct versus effort as a student perception. *Educational Psychology Review*, 32, 1163-1175. <https://doi.org/10.1007/s10648-020-09577-3>
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41(10), 1040-1048. <https://doi.org/10.1037//0003-066x.41.10.1040>
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.
- ElSayary, A. (2023). An investigation of teachers' perceptions of using ChatGPT as a supporting tool for teaching and learning in the digital era. *Journal of Computer Assisted Learning*, 40(3), 931-945. Portico. <https://doi.org/10.1111/jcal.12926>
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K., & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1199058>

- Escalante, J., Pack, A. & Barrett, A. (2023). AI-generated feedback on writing: insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20, 57. <https://doi.org/10.1186/s41239-023-00425-2>
- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, 6, 100198. <https://doi.org/10.1016/j.caeai.2023.100198>
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, 461-468. <https://doi.org/10.1016/j.chb.2017.05.045>
- Fryer, L. K., Thompson, A., Nakao, K., Howarth, M., & Gallacher, A. (2020). Supporting self-efficacy beliefs and interest as educational inputs and outcomes: Framing AI and Human partnered task experiences. *Learning and Individual Differences*, 80, 101850. <https://doi.org/10.1016/j.lindif.2020.101850>
- Gattupalli, S., Maloy, R. W., & Edwards, S. (2023). Comparing Teacher-Written and AI-Generated Math Problem Solving Strategies for Elementary School Students: Implications for Classroom Learning. *College of Education Working Papers and Reports Series*. 5. <https://doi.org/10.7275/8sgx-xj08>
- Gupta, & Chen, Y. (2022). Supporting Inclusive Learning Using Chatbots? A Chatbot-Led Interview Study. *Journal of Information Systems Education*, 33(1), 98-108. <https://jise.org/Volume33/n1/JISE2022v33n1pp98-108.html>
- Haq, M. R., & Rubya, S. (2023). An overview of chatbot-based mobile mental health apps: insights from app description and user reviews. *JMIR mHealth and uHealth*, 11(1), e44838. <https://doi.org/10.2196/44838>
- Harter, S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology*, 17(3), 300-312. <https://doi.org/10.1037//0012-1649.17.3.300>
- Hasanein, A. M., & Sobaih, A. E. E. (2023). Drivers and Consequences of ChatGPT Use in Higher Education: Key Stakeholder Perspectives. *European Journal of Investigation in Health, Psychology and Education*, 13(11), 2599-2614. <https://doi.org/10.3390/ejihpe13110181>
- Hassija, V., Chamola, V., Mahapatra, A. et al. Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognition and Computing*, 16, 45-74 (2024). <https://doi.org/10.1007/s12559-023-10179-8>
- Hawanti, S., & Zubaydullovna, K. M. (2023). AI chatbot-based learning: alleviating students' anxiety in english writing classroom. *Bulletin of Social Informatics Theory and Application*, 7(2), 182-192. <https://doi.org/10.31763/businta.v7i2.659>
- Hayat, A. A., Shateri, K., Amini, M., & Shokrpour, N. (2020). Relationships between academic self-efficacy, learning-related emotions, and metacognitive learning strategies with academic performance in medical students: a structural equation model. *BMC Medical Education*, 20, 1-11. <https://doi.org/10.1186/s12909-020-01995-9>
- Hmoud, M., Swait, H., Hamad, N., Karram, O., & Daher, W. (2024). Higher Education Students' Task Motivation in the Generative Artificial Intelligence Context: The Case of ChatGPT. *Information*, 15(1), 33. <https://doi.org/10.3390/info15010033>
- Holmes, W. (2023). *The unintended consequences of artificial intelligence and education*. Education International.
- Hsu, Huang, H.-L., Hwang, G.-J., & Chen, M.-S. (2023). Effects of Incorporating an Expert Decision-making Mechanism into Chatbots on Students' Achievement, Enjoyment, and Anxiety. *Educational Technology & Society*, 26(1), 218-231. [https://doi.org/10.30191/ETS.202301_26\(1\).0016](https://doi.org/10.30191/ETS.202301_26(1).0016)
- Impact Research. (2024). AI Chatbots in Schools: Findings from a Poll of K-12 Teachers, Students, Parents, and College Undergraduates. *Walton Family Foundation*. Accessible at: <https://8ce82b94a8c4fdc3ea6d-b1d233e3bc-3cb10858bea65ff05e18f2.ssl.cf2.rackcdn.com/bf/24/cd3646584af89e7c668c7705a006/deck-impact-analysis-national-schools-tech-tracker-may-2024-1.pdf>

- Karumbaiah, S., Ocumpaugh, J., & Baker, R. S. (2021). Context Matters: Differing Implications of Motivation and Help-Seeking in Educational Technology. *International Journal of Artificial Intelligence in Education*, 32(3), 685–724. <https://doi.org/10.1007/s40593-021-00272-0>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., & 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>
- Kaya, S., & Karakoc, D. (2022). Math Mindsets and Academic Grit: How are They Related to Primary Math Achievement? *European Journal of Science and Mathematics Education*, 10(3), 298–309. <https://doi.org/10.30935/scimath/11881>
- Khan, S. (2024). *Brave New Words: How AI Will Revolutionize Education (and Why That's a Good Thing)*. Penguin Books.
- Khurma, O. A., Albahti, F., Ali, N., & Bustanji, A. (2024). AI ChatGPT and student engagement: Unraveling dimensions through PRISMA analysis for enhanced learning experiences. *Contemporary Educational Technology*, 16(2), ep503. <https://doi.org/10.30935/cedtech/14334>
- Kikerpill, K., & Siibak, A. (2023). App-hazard disruption: An empirical investigation of media discourses on ChatGPT in educational contexts. *Computers in the Schools*, 40(4), 334–355. <https://doi.org/10.1080/07380569.2023.2244941>
- Kleitman, S., Jackson, S. A., Zhang, L. M., Blanchard, M. D., Rizvandi, N. B., & Aidman, E. (2022). Applying Evidence-Centered Design to Measure Psychological Resilience: The Development and Preliminary Validation of a Novel Simulation-Based Assessment Methodology. *Frontiers in Psychology*, 12, 717568. <https://doi.org/10.3389/fpsyg.2021.717568>
- 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>
- Kreijkes, P., Kewenig, V., Kuvalja, M., Lee, M., Vitello, S., Hofman, J., Sellen, A., Rintel, S., Goldstein, D., Rothschild, D., Tankelevitch, L., & Oates, T. (2025). Effects of LLM Use and Note-Taking On Reading Comprehension and Memory: A Randomised Experiment in Secondary Schools. Accessible at: <https://ssrn.com/abstract=5095149>. <http://dx.doi.org/10.2139/ssrn.5095149>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2022). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Kumar, H., Rothschild, D. M., Goldstein, D. G., & Hofman, J. M. (2023). Math Education with Large Language Models: Peril or Promise?. Accessible at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4641653
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00426-1>
- Lee, H. Y., Chen, P. H., Wang, W. S., Huang, Y. M., & Wu, T. T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21(1), 1–28. <https://doi.org/10.1186/s41239-024-00447-4>
- Lee, Y. F., Hwang, G. J., & Chen, P. Y. (2022). Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development*, 70(5), 1843–1865. <https://doi.org/10.1007/s11423-022-10142-8>
- Lehmann, M., Cornelius, P. B., & Sting, F. J. (2024). AI Meets the Classroom: When Does ChatGPT Harm Learning?. *arXiv preprint arXiv:2409.09047*. <https://doi.org/10.2139/ssrn.4941259>
- Li-Fang, Z. (2006). A Research on Relations among the Academic Self-concept, Achievement Goal Orientation and Learning Persistence of College Students. *Chinese Journal of Clinical Psychology*, 14(2), 192–193. Accessible at: <https://scispace.com/papers/a-research-on-relations-among-the-academic-self-concept-1et03l0heq>

- Lodge, J., Howard, S., Bearman, M., Dawson, P., & Associates (2023). *Assessment Reform for The Age of Artificial Intelligence*. Tertiary Education Quality and Standards Agency. Accessible at : <https://www.teqsa.gov.au/sites/default/files/2023-09/assessment-reform-age-artificial-intelligence-discussion-paper.pdf>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed. An argument for AI in Education*, 18. Accessible at : https://www.researchgate.net/profile/Rosemary-Luckin/publication/299561597_Intelligence_Unleashed_An_argument_for_AI_in_Education/links/5812782608ae1f5510c2aa4d/Intelligence-Unleashed-An-argument-for-AI-in-Education.pdf
- Luo, W., He, H., Liu, J., Berson, I. R., Berson, M. J., Zhou, Y., & Li, H. (2024). Aladdin's Genie or Pandora's Box for early childhood education? Experts chat on the roles, challenges, and developments of ChatGPT. *Early Education and Development*, 35(1), 96-113. <https://doi.org/10.1080/10409289.2023.2214181>
- Ma, Q. (2021). The role of teacher autonomy support on students' academic engagement and resilience. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.778581>
- Marton, F., & Saljo, R. (1976). On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, 46(1), 4-11. <https://doi.org/10.1111/j.2044-8279.1976.tb02980.x>
- Mei, Q., Xie, Y., Yuan, W., & Jackson, M. O. (2024). A Turing test of whether AI chatbots are behaviorally similar to humans. *Proceedings of the National Academy of Sciences*, 121(9). <https://doi.org/10.1073/pnas.2313925121>
- Mercer, N., Wegerif, R., & Major, L. C. (Eds.). (2019). *The Routledge international handbook of research on dialogic education*. Routledge Meyer, Turner & Spencer.
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: A guidance for policymakers*. UNESCO Publishing. Accessible at : https://books.google.co.uk/books?hl=en&lr=&id=yyE7EAAAQBAJ&oi=fnd&pg=PA3&dq=UNESCO+AI+and+Education+Guidance+for+Policymakers&ots=cBibH7riP5&sig=Laoqq_Xf87A-TaQAfR3BhymVAwc&redir_esc=y#v=onepage&q=UNESCO%20AI%20and%20Education%20Guidance%20for%20Policymakers&f=false
- Mierke, K., Scheidtmann, S., & Ibrahimova, D. (2017). Task-order autonomy fosters flow, self-efficacy, cognitive performance, and challenge-seeking in test situations. *Journal of Business and Media Psychology*, 8, 20-27. Accessible at: https://journal-bmp.de/wp-content/uploads/3_set_them_free.pdf
- Mintz, J., Holmes, W., Liu, L., & Perez-Ortiz, M. (2023). Artificial Intelligence and K-12 Education: Possibilities, Pedagogies and Risks. *Computers in the Schools*, 40(4), 325-333. <https://doi.org/10.1080/07380569.2023.2279870>
- Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S., ... & Hui, P. (2024). ChatGPT in education: A blessing or a curse? A qualitative study exploring early adopters' utilization and perceptions. *Computers in Human Behavior: Artificial Humans*, 2(1), 100027. <https://doi.org/10.1016/j.chbah.2023.100027>
- Mouratidis, A., Michou, A., Demircioğlu, A. N., & Sayil, M. (2018). Different goals, different pathways to success: Performance-approach goals as direct and mastery-approach goals as indirect predictors of grades in mathematics. *Learning and Individual Differences*, 61, 127-135. <https://doi.org/10.1016/j.lindif.2017.11.017>
- National 4-H Council & Hart Research. (2024). *Youth AI use & understanding survey*. Accessible at : https://www.4-h.org/wp-content/uploads/2024/02/27162629/Hart-Research-Youth-AI-Survey-Results42.pdf?_ga=2.151063673.1761332902.1724919083-558076093.1724919083
- Ng, D.T.K., Leung, J.K.L., Su, J. et al. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Education Technology Research and Development*, 71, 137-161. <https://doi.org/10.1007/s11423-023-10203-6>
- Novianti, I. (2025). An exploratory study on the challenges of AI technology in education and its practical recommendations. *International Journal of Social Science Research and Review*, 8(3), 26-37. <https://doi.org/10.47814/ijssrr.v8i3.2481>
- OECD (2019), *PISA 2018 Results (Volume II): Where All Students Can Succeed*, PISA, OECD Publishing. Accessible at: <https://doi.org/10.1787/b5fd1b8f-en>.

- Phung, T., Pădurean, V.-A., Cambronero, J., Gulwani, S., Kohn, T., Majumdar, R., Singla, A., & Soares, G. (2023). Generative AI for Programming Education: Benchmarking ChatGPT, GPT-4, and Human Tutors. *Proceedings of the 2023 ACM Conference on International Computing Education Research - Volume 2*, 41–42. <https://doi.org/10.1145/3568812.3603476>
- Polyportis, A. (2024). A longitudinal study on artificial intelligence adoption: understanding the drivers of ChatGPT usage behavior change in higher education. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1324398>
- Porter, T., Molina, D. C., Blackwell, L., Roberts, S., Quirk, A., Duckworth, A. L., & Trzesniewski, K. (2020). Measuring Mastery Behaviours at Scale: The Persistence, Effort, Resilience, and Challenge-Seeking (PERC) Task. *Journal of Learning Analytics*, 7(1), 5–18. <https://doi.org/10.18608/jla.2020.71.2>
- Pratschke, B. M. (2024). *Generative AI and education: digital pedagogies, teaching innovation and learning design*. Springer.
- Sharples, M. (2023). Towards social generative AI for education: theory, practices and ethics. *Learning: Research and Practice*, 9(2), 159–167. <https://doi.org/10.1080/23735082.2023.2261131>
- Shieh, E., Vassel, F. M., Sugimoto, C., & Monroe-White, T. (2024). Laissez-Faire Harms: Algorithmic Biases in Generative Language Models. *arXiv preprint arXiv:2404.07475*. Accessible at: <https://arxiv.org/abs/2404.07475>
- Singh, M. (2024). Byju's founder says his edtech startup, once worth \$22B, is now 'worth zero'. *TechCrunch*. Accessible at: <https://techcrunch.com/2024/10/17/byjus-founder-says-his-edtech-startup-once-worth-22b-is-now-worth-zero/>
- Skinner, B. F. (1953). *Science and Human Behavior*. Macmillan.
- Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips. *Science*, 333(6043), 776–778. <https://doi.org/10.1126/science.1207745>
- Strachan, J. W. A., Albergo, D., Borghini, G., Pansardi, O., Scaliti, E., Gupta, S., Saxena, K., Rufo, A., Panzeri, S., Manzi, G., Graziano, M. S. A., & Becchio, C. (2024). Testing theory of mind in large language models and humans. *Nature Human Behaviour*, 8(7), 1285–1295. <https://doi.org/10.1038/s41562-024-01882-z>
- Sufyan, N. S., Fadhel, F. H., Alkhathami, S. S., & Mukhadi, J. Y. (2024). Artificial intelligence and social intelligence: preliminary comparison study between AI models and psychologists. *Frontiers in Psychology*, 15, 1353022.
- Sulik, M. J., Finch, J. E., & Obradović, J. (2020). Moving beyond executive functions: Challenge preference as a predictor of academic achievement in elementary school. *Journal of Experimental Child Psychology*, 198, 104883.
- Sun, D., Boudouaia, A., Zhu, C. et al. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education*, 21, 14. <https://doi.org/10.1186/s41239-024-00446-5>
- Terblanche, Molyn, J., Williams, K., & Maritz, J. (2023). Performance matters: students' perceptions of Artificial Intelligence Coach adoption factors. *Coaching : an International Journal of Theory, Research & Practice*, 16(1), 100–114. <https://doi.org/10.1080/17521882.2022.2094278>
- Tinto, V. (2022). Exploring the Character of Student Persistence in Higher Education: The Impact of Perception, Motivation, and Engagement. In: Reschly, A.L., Christenson, S.L. (eds). *Handbook of Research on Student Engagement*. Springer. https://doi.org/10.1007/978-3-031-07853-8_17
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1). <https://doi.org/10.1186/s40561-023-00237-x>
- Tuomi, I. (2019). The Impact of Artificial Intelligence on Learning, Teaching, and Education: Policies for the Future. *JRC Science for Policy Report. European Commission*. Accessible at: <https://publications.jrc.ec.europa.eu/repository/handle/JRC113226>

- Vasconcelos, M. A. R., & dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2296. <https://doi.org/10.29333/ejmste/13313>
- Vassel, F.-M., Shieh, E., Sugimoto, C. R., & Monroe-White, T. (2024). The Psychosocial Impacts of Generative AI Harms. *Proceedings of the AAAI Symposium Series*, 3(1), 440–447. <https://doi.org/10.1609/aaais.v3i1.31251>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press.
- Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21(1). <https://doi.org/10.1186/s41239-024-00448-3>
- Watson, S. (2025). *Emergent discourses in generative AI in education and society* (Pre-review). CUP <https://rgdoi.net/10.13140/RG.2.2.29773.14562>
- Wegerif, R. (2019). Dialogic education. In *Oxford Research Encyclopedia of Education*. Accessible at: <https://oxfordre.com/education/display/10.1093/acrefore/9780190264093.001.0001/acrefore-9780190264093-e-396>
- Wegerif, R., & Major, L. (2024). *The theory of educational technology: towards a dialogic foundation for design*. Routledge.
- Wu, T.-T., Lee, H.-Y., Li, P.-H., Huang, C.-N., & Huang, Y.-M. (2023). Promoting Self-Regulation Progress and Knowledge Construction in Blended Learning via ChatGPT-Based Learning Aid. *Journal of Educational Computing Research*, 61(8), 3–31. <https://doi.org/10.1177/07356331231191125>
- Xenofontos, C., & Mouroutsou, S. (2023). “Resilience” in Mathematics Education Research: A Systematic Review of Empirical Studies. *Scandinavian Journal of Educational Research*, 67(7), 1041–1055. <https://doi-org.ezp.lib.cam.ac.uk/10.1080/00313831.2022.2115132>
- Xu, X., Wang, X., Zhang, Y., & Zheng, R. (2024). Applying ChatGPT to tackle the side effects of personal learning environments from learner and learning perspective: An interview of experts in higher education. *PLOS ONE*, 19(1), e0295646. <https://doi.org/10.1371/journal.pone.0295646>
- Yang, S.-H., Chen, Y.-E., Wang, J.-H., & Chen, G.-D. (2023). Integrating a Chatbot and the Concepts of Grit and Growth Mindset into a Mobile Game-Based Learning System. *Innovative Technologies and Learning*, 187–196. https://doi.org/10.1007/978-3-031-40113-8_18
- Yildiz Durak, H. (2023). Conversational agent-based guidance: examining the effect of chatbot usage frequency and satisfaction on visual design self-efficacy, engagement, satisfaction, and learner autonomy. *Education and Information Technologies*, 28(1), 471–488. <https://doi.org/10.1007/s10639-022-11149-7>
- Yilmaz, R. & Karaoglan Yilmaz, F.G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students’ computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, 4, 100147. <https://doi.org/10.1016/j.caeai.2023.100147>
- Yin, M., Heard, H. E., Szabo, J., & Ankoor, N. (2021). *Social and Emotional Skills of 15-Year-Old Students in the Houston Independent School District*. Houston Education Research Consortium. Accessible at: <https://files.eric.ed.gov/fulltext/ED617735.pdf>
- Young, J. R. (2024). An Education Chatbot Company Collapsed. Where Did the Student Data Go? In *EdSurge*. Accessible at: <https://www.edsurge.com/news/2024-07-15-an-education-chatbot-company-collapsed-where-did-the-student-data-go>
- Zamfiroiu, A., Vasile, D., & Savu, D. (2023). ChatGPT—A Systematic Review of Published Research Papers. *Informatica Economica*, 27(1), 5–16. <https://doi.org/10.24818/issn14531305/27.1.2023.01>
- Zhu, C., Sun, M., Luo, J., Li, T., & Wang, M. (2023). How to harness the potential of ChatGPT in education?. *Knowledge Management & E-Learning*, 15(2), 133–152. <https://doi.org/10.34105/j.kmel.2023.15.008>

Acknowledgements

This work was funded by the Cambridge Trust, Hughes Hall, Accelerate Science Programme for Scientific Discovery, and the Cambridge Centre for Data-Driven Discovery.

About the authors



Megan Ennion

Megan is a PhD researcher at the University of Cambridge focusing on the psychological and behavioural impacts of Generative AI. She also leads on several studies exploring the use and attitudes towards AI use amongst students and teaching staff in England and is working on a project exploring worldwide university policy towards AI in her role as research assistant at the Digital Education Futures Initiative (DEFI) based at Hughes Hall, University of Cambridge. Megan is the founder of the AI & Education Community, a collective initiative of staff and students at the University of Cambridge interested in or working on AI and education. She is particularly interested in how AI can be leveraged to improve educational experiences and outcomes.

✉ mle34@cam.ac.uk

ORCID [0009-0008-1899-1523](https://orcid.org/0009-0008-1899-1523)



Ros McLellan

Ros McLellan is Associate Professor in the Faculty of Education, University of Cambridge, currently in the role of Director of Learning & Teaching. She is a social psychologist with a background in secondary teaching in the UK. Her research interests lie in the fields of student wellbeing and motivation, and she enjoys working with practitioners seeking to support student wellbeing and motivation. Recent research as PI includes collaborative work to investigate young people's wellbeing in Kazakhstan, an examination of the impact of creative practitioners working with young people in the classroom, and transition from primary to secondary school. She was also a CI on a project on mathematics anxiety. She is currently working with a charity on a theory of change based on self-efficacy. She co-edited a volume in the EERA transdisciplinary perspectives in educational research series on wellbeing and schooling bringing together cross cultural and cross disciplinary perspectives published last year. She is Chair of BERA's Publications Committee, and a member of the Mental Health and Wellbeing SIG.

✉ rwm11@cam.ac.uk

ORCID [0000-0002-7583-5550](https://orcid.org/0000-0002-7583-5550)

Open Access (CC BY 4.0)

© 2025 The Authors. This article is distributed under Creative Commons Attribution 4.0 International licence.



You are free to

- Share — copy and redistribute the material in any medium or format
- Adapt — remix, transform, and build upon the material for any purpose, even commercially.

Under the following terms:

- Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

The full licence conditions are available at: <https://creativecommons.org/licenses/by/4.0/>

