MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2025, 21(10), em2713
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/17077

Integrating generative artificial intelligence chatbots into chemistry teaching: Impact of affective factors on engagement and conceptual understanding

Samuel Jere [1*], Mamotena Mpeta [1]

[1] University of Venda, Thohoyandou, SOUTH AFRICA

Received 30 June 2025 ▪ Accepted 27 August 2025

Abstract

There has been a recent interest in leveraging generative artificial intelligence, large language models, to assist secondary school learners in improving their engagement and conceptual understanding (CU) of challenging concepts in chemistry. However, most of these studies have focused only on academic achievement. The influence of affective factors when integrating large language models has largely been ignored. The study investigated the effects of self-efficacy (SE), motivation and satisfaction on learner engagement and CU when ChatGPT was integrated into chemistry lessons. The self-regulatory learning (SRL) model was the theoretical framework used for the study. A cross-sectional survey design was employed in this quantitative study. Two schools in the Vhembe East District of Limpopo in South Africa participated in the study. A questionnaire was used in data collection after learners were exposed to intermolecular forces in physical sciences, and ChatGPT was integrated into their lessons. The sample size from the two schools was 240 learners. Structural equation modelling and path analysis were used to analyse the data. The study revealed that satisfaction significantly enhanced both engagement and perceived CU. In contrast, SE impacted perceived CU only, while motivation solely improved engagement. The study has implications for teachers integrating artificial intelligence tools like ChatGPT in teaching chemistry. The findings extend our understanding of the practical implications of the SRL model when integrating ChatGPT into instructional practices.

Keywords: artificial intelligence, ChatGPT, physical sciences, self-regulatory learning model, self-efficacy, motivation, satisfaction, learner engagement, conceptual understanding

INTRODUCTION

ChatGPT was developed and launched by OpenAI in November 2022, bringing massive potential to the transformation of the way children learn (Lee et al., 2024). ChatGPT is a large language model capable of performing complex tasks such as providing an environment for personalised and interactive learning, creating formative assessment tasks, and giving learners feedback (Baidoo-Anu & Ansah, 2023; Jere et al., 2024). Artificial intelligence (AI) tools can serve as teaching assistants or tutoring agents when integrated into teaching. Due to this ability, they have been described as 'proxy teachers' capable of assisting learners on behalf of the teacher (Chiu et al., 2023). Researchers recognise the potential of AI tools such as ChatGPT to contribute significantly to enhancing teaching practices (Cooper, 2023; Gill et al., 2024).

ChatGPT was developed based on generative AI, which creates artificial content using existing digital resources (Baidoo-Anu & Ansah, 2023). Generative AI has two forms: generative adversarial network (GAN) and generative pre-trained transformer (GPT). GAN is used mainly for voice generation, graphics, and videos, while GPT is used for natural language processing and text production. GPT is a language model based on a transformer architecture (Kasneci et al., 2023). It is pre-trained in a generative and unsupervised way. It evolved from GPT-2 to GPT-3; the current version is GPT-4. GPT-3 had 175 billion parameters and 499 billion words, costing a staggering 4.6 million United States dollars to train. The capabilities of GPT-3 include the generation, classification, and summarisation of text. It can also

✉ samuel.jere@univen.ac.za **(*Correspondence)** ✉ Mamotena.Mpeta@univen.ac.za

Contribution to the literature

- Although there is growing interest in leveraging AI to enhance learners' conceptual understanding, empirical evidence on how affective factors impact learning when using AI remains limited
- Thus, this study contributes to an understanding of how affective factors influence learning while integrating AI in chemistry education.
- Effective integration of AI in chemistrv education is not just a technological issue, but it is also a pedagogical matter that requires consideration of affective factors.

translate language, answer questions, and recognise entities. GPT-4 generates more accurate and quality responses to users' prompts than earlier versions (Küchemann et al., 2023).

Generative AI technologies can be integrated into teaching chemistry, bringing many benefits to learners and teachers. For example, in formative assessment, generative AI technology can be used to generate open-ended or structured practice questions, multiple choice questions and other forms of quizzes used to help learners better understand and retain the content they would be studying (Kasneci et al., 2023). Chemistry requires learners to be adept at problem-solving, and language models like ChatGPT can be instrumental in this regard (Jere, 2025). They can be used to offer step-by-step explanations of how to solve complex chemistry problems, thus helping learners' conceptual understanding (CU) of the reasoning required to solve these difficult problems (Kasneci et al., 2023). Furthermore, generative AI can be used by the learners to generate similar practice problems, and the AI can then be used to assess the learners' responses, providing learners with clarity on aspects they find challenging.

While integrating generative AI in teaching, learners must be guided in learning how to produce specific prompts that can elicit a suitable response from the AI chatbot, as it was noted that the quality of the response the user obtains depends on the quality of the prompt (Kıyak, 2023). Chiu et al. (2023) found that teacher assistance in prompt generation for novice learners improves motivation when integrating AI tools into instruction. The development of appropriate questions that yield desired responses to the learner's needs has been referred to as prompt engineering (Jacobsen & Weber, 2023). Teachers would need to develop the learner's ability to generate prompts that are clear, logical, explicit, adaptive and reflective (Lo, 2023). The quality of the feedback from ChatGPT depends mainly on the specificity and clarity of the prompt (Jacobsen & Weber, 2023). Hence, during the integration of ChatGPT in lessons, the teacher should guide the learner in developing self-efficacy (SE) to produce appropriate prompts.

When integrating generative AI, such as ChatGPT, into teaching chemistry, it is essential to investigate affective factors influencing learning outcomes. AI technologies offer personalised learning environments that can positively impact learner engagement and CU. If affective factors such as attitudes towards learning, the feeling of being cared for and valued, happiness, well-being and satisfaction are improved through lessons that integrate AI (Nguyen et al., 2022), it can result in greater CU.

Rationale and Research Questions for the Study

Previous studies have revealed that integrating some kinds of digital technologies, such as simulations, impacts affective factors such as learner motivation (LM), satisfaction and SE. These factors have also been shown to influence CU and engagement. While numerous studies have explored the effects of AI chatbots on assessment and evaluation in the broader field of science education (Almasri, 2024; Cooper, 2023), there is a notable scarcity of empirical studies that investigate the effect of affective factors on learning outcomes when integrating these emerging technologies into chemistry education.

Generative AI tools such as ChatGPT are emerging digital technologies, and there is a paucity of research on how integrating these tools influences affective factors, nor how these affective factors play a role in the learners' CU and engagement. Therefore, this study investigated the impact of SE, motivation and satisfaction on the learners' CU and engagement when ChatGPT is integrated into physical sciences teaching. The study's research question was: How do SE, motivation and satisfaction influence CU and learner engagement when integrating ChatGPT into physical sciences teaching? The study was guided by Zimmerman's (2000) self-regulatory learning (SRL) model as the theoretical framework to study how SE, motivation and satisfaction influenced CU and engagement while integrating ChatGPT in learning chemistry.

LITERATURE REVIEW

Self-Regulatory Learning Model

SRL refers to processes and beliefs that allow learners to transform their mental abilities into academic skills (Zimmerman, 2008). SRL is a prerequisite for effective technology-enhanced learning (Chiu et al., 2023). When integrating AI into teaching, learners must self-regulate to understand the learning tasks. While teaching complex topics in chemistry using AI, learners would

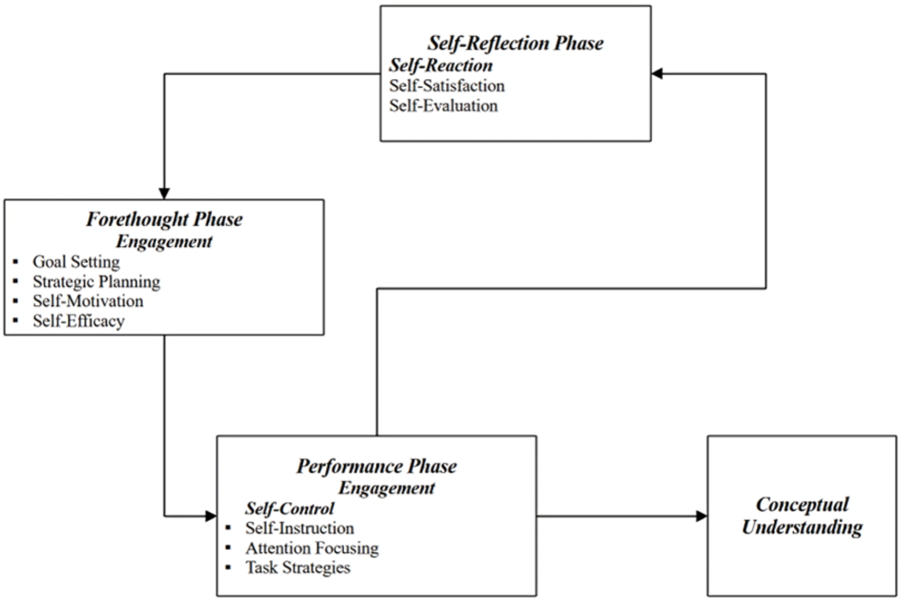

Figure 1. Phases of SRL (adapted from Zimmerman, 2008)

benefit from being guided in the three phases of SRL proposed by Zimmerman (2008). The three phases are forethought, performance and self-reflection and form a cyclical process involving activities that occur before, during and after the learning task (**Figure 1**).

The forethought phase involves task analysis, setting goals and means by which these goals can be achieved (Wu et al., 2024). A strong belief in their ability to successfully perform the task with the help of the model raises their SE (Bandura et al., 1999). They also need self-motivation to be able to perform the task well. The forethought phase occurs in this study when the learners are provided with the learning task as they analyse it and begin engaging with the task. This is immediately followed by the performance phase, when the learner engages with the task (**Figure 1**).

During the performance phase, learners engage with the learning task, and learning occurs through modelling, during which the learner compares their thoughts, beliefs, and behaviours against a model's (Schunk & Zimmerman, 2007). In the integration of AI using ChatGPT, learners can use the chatbot as a model for observational learning. This enables users to reflect on their thoughts and beliefs while interacting with the chatbot (Schunk & Zimmerman, 2007).

In the engagement phase, learners must attend to the chatbot's information and decide to seek further clarity from the model. By concisely phrasing their prompts, ChatGPT can offer learners precise answers to their enquiries (Wu et al., 2024). Learners mentally code and transform modelled information from the chatbot and rehearse the information (Schunk & Zimmerman, 2007). Motivation is important as the learners gain greater CU to retain the information they are learning. Through observational learning, they learn the sequence of action needed in problem-solving from the model (Schunk & Zimmerman, 2007), further strengthening their SE in solving chemistry problems. This may lead to CU. The final phase is self-reflection. This is achieved through self-evaluation and judgement of their ability to perform the task (Wu et al., 2024). This may lead to self-satisfaction if they have mastered the learning task. The SRL model indicates that affective factors such as motivation, SE, and satisfaction are critical to observational learning to attain important educational goals, such as CU, when engaging with the learning task. The integration of ChatGPT in chemistry learning depends mainly on the learner's ability to pose relevant questions to the chatbot to enhance SRL.

Integrating ChatGPT into Learning Physical Sciences

When integrating ChatGPT into learning chemistry, learners interact with the machine interface through questioning. Learners will either ask ChatGPT to generate questions for them to practice or ask the chatbot to answer questions. Questioning engages learners in active retrieval and practice of what they are learning, and this leads to meaningful learning of the concepts, enhancing their ability to apply the learned concepts in novel situations (Karpicke, 2012; Karpicke & Grimaldi, 2012). In Bloom's taxonomy, questions that require analysis, synthesis and evaluation skills are considered high-level questions (Al Faraby et al., 2023) while those requiring recall of information in chemistry are low-level questions. Constructing high-level questions is time-consuming and cognitively demanding for teachers. The advent of deep learning models in AI has shown great promise in automatic question generation for high-level questions (Al Faraby et al., 2023). In this regard, ChatGPT can be integrated into chemistry learning by prompting it to generate questions for learners to practice.

AI tools, such as ChatGPT, have the potential to positively impact educational outcomes in learning chemistry. Few studies, however, have explored the integration of AI technology in education, and currently, there is no clear understanding of how these tools influence important educational goals (Chiu et al., 2023). Some of the few studies on the use of AI in education have indicated that AI chatbot use results in enhanced academic performance (Xu & Ouyang, 2022). However, other studies have revealed that integrating AI does not affect learners' critical thinking skills (Jia & Tu, 2024). This then calls for further studies to determine if AI chatbots can improve the learners' CU.

Impact of Motivation on Conceptual Understanding and Engagement

Motivation refers to psychological factors such as interest and mental efforts learners exert to achieve some educational goals (Li & Li, 2022). Motivation is thought to play a critical role in influencing learner engagement. In a quasi-experimental study in which learners in the experimental group used a generative AI chatbot (ChatGPT) and the control group used standard search engines in blended learning of mathematics, Wu et al. (2024) found that learning using the chatbot increases the learners' engagement. They concluded that using chatbots in blended learning enhances learners' self-regulation. Similarly, Qawaqneh et al. (2023) investigated the influence of AI-based virtual experiments on motivation at the grade seven level. They found that AI learning environments enhanced learners' motivation and CU of mathematics.

Iyamuremye and Ndihokubwayo (2024) investigated learners' understanding of atomic structure and chemical bonding after integrating ChatGPT into science teaching. They found that integrating ChatGPT increased academic performance by 16.6% and enhanced the learners' engagement and motivation. These results are supported by dos Santos (2023), who found that AI chatbots such as ChatGPT provide accurate, comprehensive and detailed responses to chemistry questions and can assist learners in developing diverse skills such as critical thinking, problem-solving and creativity. dos Santos (2023) argues that educators can leverage AI chatbots to overcome learner disengagement in learning chemistry.

Based on the reviewed studies in the previous paragraphs and the theoretical framework, we propose the following hypotheses:

H1. The learners' motivation positively and significantly impacts learner engagement when integrating ChatGPT into learning.

H2. The learners' motivation significantly positively influences CU when integrating AI technology into learning physical sciences.

Influence of Self-efficacy on Conceptual Understanding and Engagement

In integrating generative AI tools in learning, the learners' SE beliefs are critical in achieving desired outcomes such as CU. SE is a person's belief in their ability to achieve desired outcomes despite obstacles (Bandura, 2012). SE beliefs influence motivation and learners' actions, and thus, it is pivotal to achieving learning outcomes (Bandura, 2012). In addition, the integration of AI in learning was empirically found to have a positive and significant effect on knowledge construction in mathematics compared to those not using AI (Wu et al., 2024). As knowledge construction should involve CU, we can deduce that AI technology can lead to increased CU. Therefore, it can be hypothesised that the learners' SE beliefs directly influence their CU when learning using generative AI tools. Furthermore, their SE beliefs may impact their engagement with the learning tasks.

When learning science, engagement has been defined as the extent to which learners actively and productively participate in an activity (Ben-Eliyahu et al., 2018). If learners have high SE, it can lead to increased attempts to overcome obstacles that negatively impede their engagement. Furthermore, an empirical study by Ben-Eliyahu et al. (2018) found that SE was positively related to engagement when doing science activities. Based on the reviewed previous studies and the self-regulatory theoretical framework, the following hypotheses are proposed:

H3. SE has a direct and significant positive impact on learners' engagement.

H4. SE significantly and positively affects learners' CU when integrating generative AI tools in learning physical sciences.

Effect of Satisfaction on Conceptual Understanding and Learner Engagement

Long (1985) defines learner satisfaction (LS) as a pleasant feeling or attitude towards learning experiences. Some learning experiences, such as blended learning utilising digital technology, have enhanced LS (Hua et al., 2024). Similarly, Du (2023) studied the factors influencing learners' satisfaction in massive open online courses. The findings indicate that teaching strategies such as the use of video during instruction, the nature of the content and learner evaluation play a role in LS. Similarly, Rajabalee and Santally (2021) investigated the relationship between learners' satisfaction and engagement in an online course. They found a significant, positive correlation between LS and engagement. LS and engagement were also positively correlated with academic achievement, although the association was weak.

When ChatGPT is integrated into physical sciences learning, learners obtain personalised feedback to their

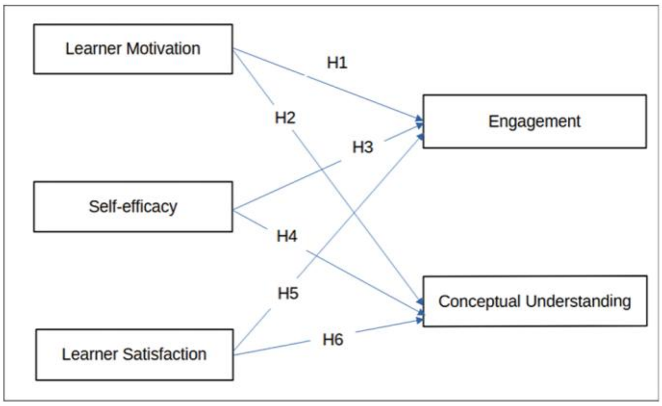

Figure 2. Theoretical model of the effects of SE, motivation, and satisfaction on engagement and CU (Source: Authors' own elaboration, using LibreOffice Draw)

inquiries, ask questions leading to immediate responses, and seek clarity on concepts that they do not understand. Such engagement with ChatGPT is likely to lead to CU. Hence, based on the reviewed studies and the theoretical framework, we propose the following hypotheses:

H5. LS positively and significantly impacts learners' engagement.

H6. LS has a significant positive effect on CU.

Influence of Affective Factors on Conceptual Understanding and Engagement

The theoretical model in **Figure 2** summarises the influence of SE, motivation and satisfaction on engagement and CU. This model guided this study and was empirically verified through structural equation modelling.

METHODOLOGY

A cross-sectional survey design was used in the study. In this design, data is collected from participants at a single point in time (Van der Stede, 2014) to provide a snapshot of the prevailing situation. The design offers advantages such as being quick to conduct, relatively inexpensive, and may have a high response rate (Spector, 2019). They enable researchers to determine whether pairs of variables are related and whether moderators are involved (Spector, 2019). By anchoring the hypotheses in theory and developing a plausible theoretical research model based on prior research, some of the limitations of cross-sectional research can be mitigated (Van der Stede, 2014).

Population and Sampling

The population of the study were all grade 11 physical sciences learners in the Vhembe East District of Limpopo. Physical sciences in South Africa consist of physics and chemistry. Two schools were randomly selected to participate in the study. All physical sciences learners in these schools participated. The researchers held workshops with the teachers of these learners to demonstrate how to integrate ChatGPT in teaching the chemistry topic matter and materials–intermolecular forces. The teachers would use their conventional teaching strategies but reserve twenty minutes to integrate ChatGPT in each lesson for two weeks.

Context

Intermolecular forces is a topic in which the learner would study the different types of van der Waal, ion-dipole, and ion-induced dipole forces. They study the relationship between molecular size and strength of intermolecular forces and explain the effect of intermolecular forces on melting and boiling points. The topic also covers thermal expansion and conductivity, density of materials and the chemistry of water. Anecdotal evidence suggests that, generally, learners find this topic challenging. To attempt to ameliorate these challenges, the teachers were asked to integrate ChatGPT into their conventional instructional practices. This was done by asking them to devote at least twenty minutes in their lessons to asking learners to use ChatGPT either to generate questions from the lesson they presented and answer those questions or to ask ChatGPT to generate answers to specific questions.

The rationale for carrying out the intervention over a period of two weeks was informed by practical and theoretical considerations. Firstly, the topic Intermolecular Forces is relatively short, as reflected in the curriculum documents in South Africa and is allocated a total of 10 hours teaching time (Department of Basic Education, 2011). Over the two-week intervention period, learners were taught the topic for a duration of 10 hours. This implies that they used ChatGPT for a total of 200 minutes. Typical of exploratory studies, the investigation sought to determine the immediate impact of the AI Chatbots as opposed to long-term gains. Similar studies, which sought to determine the impact of integrating AI on learning outcomes, were carried out in similar time frames (Lee et al., 2023; Polakova & Klimova, 2024).

An example of a question that learners would request ChatGPT to answer is provided in **Figure 3**. This is a question for an end-of-year examination. Learners would be expected to practice answering such questions, and their end-of-year exam would include such questions.

ChatGPT would provide comprehensive responses to these questions, which the teachers hoped would make it easier for learners to grasp the concepts. For example, the response from ChatGPT to question 3.1 was:

> The higher boiling point of NH_3 (-33.3 °C) compared to N_2 (-195.8 °C) is due to the difference in the intermolecular forces between these

The reaction below is used in the Haber process to manufacture ammonia.

$$N_2(g) + 3H_2(g) \rightarrow 2NH_3(g)$$

The boiling points of the substances in the reaction are as follows:

SUBSTANCE	BOILING POINT (°C)
H_2	-252,9
N_2	-195,8
NH_3	-33,3

3.1 Refer to the intermolecular forces and explain the difference in boiling point between NH_3 and N_2. (3)

3.2 Write down the FORMULA of the substance in the table that will have the lowest melting point. (1)

3.3 Explain why H_2 will evaporate faster than N_2. Refer to the type and relative strength of the intermolecular forces. (3)

3.4 Write down the FORMULA of the substance in the table that will have the highest vapour pressure. Explain your answer. (3)
[10]

Figure 3. Example of a question that ChatGPT would be prompted to answer (Department of Basic Education, 2018)

> molecules. NH_3 molecules experience **hydrogen bonding**, which is a strong intermolecular force that occurs because of the highly electronegative nitrogen atom bonded to hydrogen. This hydrogen bonding significantly increases the amount of energy required to separate the molecules, raising the boiling point. On the other hand, N_2 molecules are nonpolar and experience only weak **London dispersion forces**, which require less energy to overcome, resulting in a much lower boiling point.

When implementing the intervention, the teachers would follow the three phases of the SRL model. The forethought phase involved the teacher assisting the learners in setting the goals and explaining how they would use ChatGPT to achieve them. The performance phase involved the learners engaging with ChatGPT on their learning devices to achieve their goals. Finally, the self-reflection phase involved the learners asking themselves, *"What went well during the implementation phase?"*; *"What can be done to improve next time?"*; and *"What help do I need?"*. They kept a diary of these activities for continuity during the next lesson. The topic was taught for two weeks in the participating schools. A questionnaire was administered soon after the intervention.

Research Instrument

We developed a questionnaire from previous studies on the effects of SE, motivation, and satisfaction on learners' CU and engagement. The questions on SE were adopted from a validated questionnaire developed by Wang and Chuang (2024). The items on engagement were adopted from Wang et al. (2016). Those from LS were originally developed by Al-Momani and Pilli (2021), while items on LM were developed by the authors.

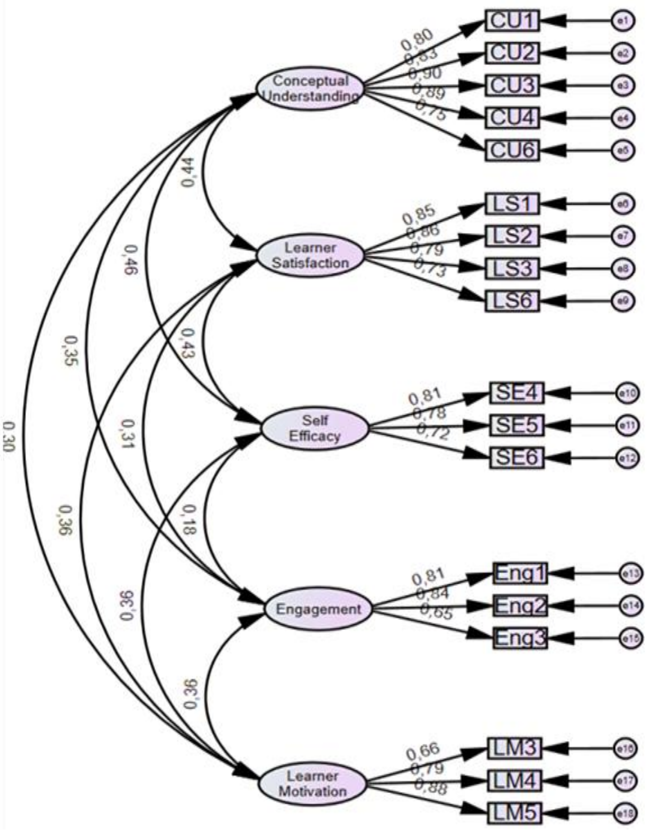

Figure 4. CFA model–Impact of affective factors on engagement and CU (Source: Authors' own elaboration, using SPSS Amos v29)

Statistical Methods in Data Analysis

Structural equation modelling (SEM), a multivariate statistical technique, using SPSS Amos version 29, was the statistical method used in data analysis (Shi & Maydeu-Olivares, 2020). In the SEM, confirmatory factor analysis (CFA) was used to evaluate the measurement model. This was followed by describing the structural model through path analysis to assess the direct effects among latent and observed variables (Shi & Maydeu-Olivares, 2020).

RESULTS

Measurement Model

CFA was performed to test the measurement model and to assess whether the observed variables (questionnaire items) closely measured the latent variables (unobserved constructs). Structural equation modelling using the maximum likelihood estimation was used to achieve this (Figure 4). The purpose was to determine the model fit and the validity and reliability of the research instrument. SPSS Amos version 29 was used in all analyses. The model consisted of three exogenous variables–LM, LS, and SE–and two endogenous variables–engagement and perceived CU.

To determine the fitness of this model, Chi-square goodness, standardised root mean square residuals

Table 1. Model fit measures (Bentler & Bonett, 1980; Hu & Bentler, 1998)

Fit measure	Recommended	Measurement model	Interpretation
RMR	< .08	.039	Excellent
GFI	> .90	.914	Acceptable
CFI	> .90	.969	Excellent
TLI	> .95	.962	Acceptable
RMSEA	< .08	.050	Excellent
AGFI	> .08	.882	Acceptable
NFI	> .90	.921	Acceptable
IFI	> .90	.969	Excellent

Table 2. Internal consistency, CR and AVE for the research instrument

Construct	Cronbach's alpha	CR	AVE
CU	0.881	0.920	0.699
LS	0.920	0.883	0.655
SE	0.809	8.814	0.594
Engagement	0.811	8.813	0.595
LM	0.813	0.823	0.611

Table 3. Fornell-Larker criterion to assess discriminant validity

	CU	LM	LS	SE	Eng
CU	**0.836**				
LM	0.301	**0.781**			
LS	0.440	0.360	**0.809**		
SE	0.460	0.355	0.430	**0.771**	
Eng	0.353	0.363	0.313	0.181	**0.771**

(SRMR), goodness of fit index (GFI), comparative fit index (CFI), relative fit index (RFI), and root mean square error of approximation (RMSEA) were extracted from Amos. Chi-square is a fit measure sensitive to sample size, and all its parameters were satisfactory except that χ^2 was significant. However, experts suggest it must be nonsignificant for a good fit (Marsh & Balla, 1994). Chi-square goodness, $\chi^2(125) = 198.687$, $\chi^2/df = 1.589$, $p < .05$, was considered a good fit as χ^2/df was less than 3 (Marsh & Balla, 1994). The SRMR, GFI, CFI, RFI, and RMSEA values in **Table 1** confirmed the model's fitness. We deduced that the model had a good fit as all values in **Table 1** were either acceptable or excellent.

The reliability of the instrument was assessed through internal consistency and composite reliability (CR). A scale with a Cronbach's alpha value of over 0.7 is considered acceptable internal consistency. As all values in **Table 2** were over 0.7, the instrument's internal consistency was acceptable (Hair et al., 2014). CR was used to ascertain the reliability of the constructs investigated. CR values above 0.7 confirm construct reliability, which was found to be the case, as shown in **Table 2** (Hair et al., 2019).

Discriminant validity, a measure of how distinct a construct is from other constructs (Sarstedt et al., 2021), was determined using the Fornell-Larker criterion.

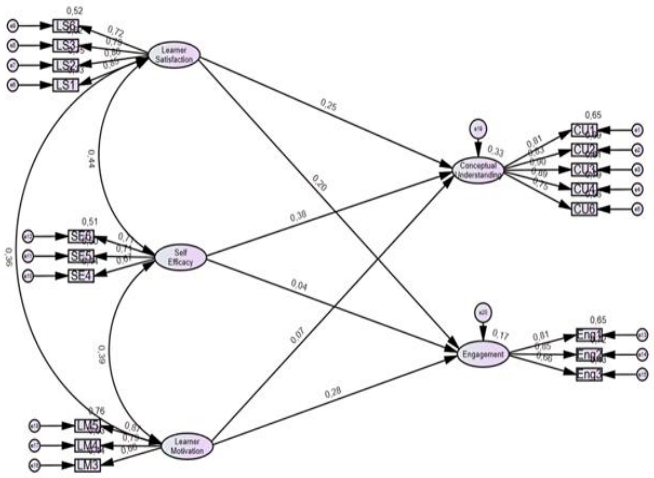

Figure 5. Structural model showing the effects of affective factors on engagement and CU (Source: Authors' own elaboration, using SPSS Amos v29)

Table 4. Path analyses and hypothesis testing

Hypothesis	Path	β	p	Decision
H1	LM → Eng	0.28	< .05	Supported
H2	LM → CU	0.07	> .05	Not supported
H3	SE → Eng	0.04	> .05	Not supported
H4	SE → CU	0.38	< .05	Supported
H5	LS → Eng	0.20	< .05	Supported
H6	LS → CU	0.25	< .05	Supported

According to Fornell and Bookstein (1982), a construct has achieved discriminant validity if the square root of its average variance extracted (AVE) is higher than its correlations with all other constructs. A visual inspection of **Table 3** suggests that all constructs achieved discriminant validity.

Structural Model

The structural equation modelling in Amos produced the structural model in **Figure 5**. The standardised path coefficients were used to examine the hypotheses in the study. The path coefficients (β) are summarised in **Table 4**.

Table 4 shows that four of the six hypotheses were supported, while two were not supported. Those not supported include SE's influence on engagement and LM's effect on CU.

Learner motivation

Hypothesis H1: The finding of the study was that the learners' motivation while integrating ChatGPT into learning has a positive, significant influence on engagement (β = 0,28, $p < 0.5$). Hence, H1 was accepted. This shows that when learners are interested and enjoy using ChatGPT in learning, they actively participate in the learning tasks, enhancing their engagement.

The engagement was most likely enhanced by the interactive and personalised experiences that learners were exposed to while integrating ChatGPT in learning intermolecular forces concepts. The increased sense of control over their learning and the teacher acting as a facilitator increased positive emotional experiences, boosting engagement.

Hypothesis H2: Path analysis in **Figure 3** shows that LM has a minor positive, non-significant effect on perceived CU (β = 0.07, p > 0.5). Based on this, **H2** was rejected. The study found no evidence that motivation has a direct significant influence on learners' CU while integrating ChatGPT when learning physical sciences. This suggests that another factor could have mediated the influence of LM on CU. As the mediating effects were not investigated in this study, no clear conclusion can be drawn as to why motivation had a limited influence on CU.

Self-efficacy

Hypothesis H3: The path coefficient showed that while integrating ChatGPT into chemistry, SE has no effect on learners' engagement (β = 0.04, p > .05). Hence, **H3** was rejected as no evidence was found that SE had an impact on learner engagement. This result suggests that SE does not directly affect engagement when learning using Chat GPT, but it does not rule out the possibility of indirect effects of SE on engagement. Some possible reasons SE was found not to affect engagement include that the ChatGPT user interface was relatively easy to use and learners could intuitively navigate through it. The presence of visual elements in the interface made it user-friendly. The teacher's assistance to learners in prompting ChatGPT could also have played a significant role. Therefore, learners did not require elevated SE levels to interact with ChatGPT.

Hypothesis H4: The findings revealed a significant, positive correlation between SE and CU (β = 0.38, p < .001). Therefore, H4 was accepted. This means learners with a strong belief in their ability to effectively use ChatGPT in learning chemistry may be expected to gain a deeper understanding of the scientific concepts. In contrast, those who doubt their ability to use ChatGPT may suffer from a lack of CU.

While interacting with ChatGPT, fear and anxiety are minimised as the chatbot is non-judgmental. Using the SRL model enhanced their CU, as setting personalised goals kept them focused on understanding challenging concepts of intermolecular forces. Furthermore, the immediate feedback from the chatbot on learners' queries creates a learning environment that fosters understanding.

Learner satisfaction

Hypothesis H5: The study found that LS was a significant positive predictor of engagement (β = 0.20, p < .05). Therefore, H5 was accepted. This result suggests that when learners are happy and satisfied with utilising ChatGPT, they are more actively involved with learning, which may lead to better learning outcomes. Some possible consequences of satisfaction with their learning experiences while integrating ChatGPT include improved emotional experiences, lower anxiety levels and increased interest in inquiry. All these factors can result in enhanced engagement, directly improving learning outcomes.

Hypothesis H6: The path coefficient demonstrates a positive and significant direct correlation between LS and perceived CU (β = 0.25, p < .05). This implies that grasping scientific concepts becomes easier when learners are happy and contented with their learning experiences using ChatGPT. The enhanced satisfaction while integrating ChatGPT in learning intermolecular forces most likely increased their curiosity and boosted exploration and inquiry. Therefore, receiving immediate feedback on their inquiries helped them overcome misconceptions, resulting in greater CU.

LS, SE and LM collectively explained 33% of the variance in the learners' CU and 17% of the variance in engagement (**Figure 5**). This implies that while these three factors are not the only determinants of CU and engagement, it is important for educators to be cognisant of them while integrating artificial intelligence chatbots into instructional practices.

DISCUSSION

Affective factors such as motivation and LS are critical in influencing learning outcomes in complex subjects such as chemistry. Therefore, it is desirable that when integrating AI, the teacher takes measures to improve variables such as motivation and LS. For example, a study by Chiu et al. (2023) reveals that intrinsic motivation and competence to learn with the chatbot depend on the learner's expertise and teacher support and influence learning outcomes.

This study has provided empirical evidence that motivation has a direct, significant impact on learner engagement. This aligns with previous studies that demonstrated that motivation positively influences learner engagement in technology-mediated learning (Yu et al., 2021). When a teaching strategy results in greater learners' motivation, their engagement with the learning task increases. This aligns with the SRL model, which proposes that the affective factors during the forethought phase can have positive effects on engagement with the learning task.

The study has shown that the learners' engagement increases when they are enthusiastic and satisfied with the learning environment utilising AI chatbots. This was also found to be the case in recent studies that demonstrated that LS and engagement are related. For example, Lane et al. (2021) investigated satisfaction and engagement in blended learning and found that engagement positively influenced satisfaction. CU and satisfaction have also been positively related in this study. This is supported by a study conducted at the university level by Abuhassna et al. (2020), who found that satisfaction with online learning results in increased academic achievement. The learners' satisfaction is enhanced when they evaluate themselves in the self-reflection phase, during which they determine what led to their success or failure in the engagement phase, as discussed in the SRL model. This satisfaction then directly impacted future forethought and engagement (Zimmerman, 2008).

Some recent studies provided empirical evidence that in learning physics and chemistry concepts, SE improves academic achievement and CU (Kalender et al., 2020; Kolil et al., 2020). This means that the learners' belief in their abilities to successfully perform tasks related to solving problems in physics and chemistry is critical in enhancing their academic achievement and understanding. If they have low self-belief, then this negatively impacts their understanding of chemistry and physics concepts, which lowers their academic achievement. The SRL model also supports the notion that the learners' affect during the forethought and performance phases is critical in their CU of the learning task, as empirically verified in this study.

Chiu et al. (2023) found that teacher support was critical in enhancing LM and satisfaction in AI-mediated learning, as the support improved the learner's expertise in using AI. The support offered by the teacher in guiding the learners in using AI helps them when they interact with the chatbots. The chatbot algorithms can evaluate the learners' level of understanding and dynamically adjust the level of difficulty of the questions, thereby enhancing the learners' understanding of the learning task (Rizvi, 2023). Therefore, it can be concluded that teacher support in this study positively impacted their affective states and improved the learning outcomes.

Implications for Practice and Future Research

The study revealed that learner-centred instruction when integrating AI large language models into chemistry teaching enhances LS, which improves engagement and CU. Furthermore, the study found that SE enhanced CU, and motivated learners showed greater engagement with learning tasks. These findings have significant implications for using self-regulatory teaching approaches where the teacher assumes the role of a facilitator. Practicing teachers should, therefore, consider integrating AI chatbots within the self-regulating teaching framework to enhance CU of challenging chemistry concepts. The study suggests that integrating AI chatbots can enhance effective teaching using self-regulatory teaching approaches. However, when integrating AI chatbots, there is a caveat that teachers should always be cognizant of.

One of the challenges of integrating chatbots, such as ChatGPT, into instructional practices is that they must be used with traditional sources of information (Limna et al., 2023). This is due to the fact that sometimes the chatbots can offer inaccurate information. When learners suspect the chatbot's information is inaccurate, they must consult the teacher or another source, such as a textbook. It would also be prudent for the teacher to determine the accuracy of the chatbots in answering specific questions during lesson preparation.

The implications of integrating self-regulatory teaching practices while using AI chatbots to assist learners in grasping challenging concepts require further research. This research area can benefit from longitudinal studies over longer timeframes. Future studies may also address the limitations inherent in this study. This was a cross-sectional study, and causal relationships cannot be inferred. Therefore, future research should consider other study designs, such as experimental or quasi-experimental approaches, to mitigate the shortcomings of the cross-sectional approach.

CONCLUSION

The study has demonstrated that affective factors enhance learner engagement and CU when learning challenging chemistry concepts while integrating AI chatbots. The SRL model has significant potential to be used as a conceptual framework when using AI chatbots, as it enhances learner-centred instructional practices. The teacher facilitates the learning process while the learners engage with the learning tasks by integrating AI chatbots. Chatbots can be integrated into teaching chemistry to take a break from traditional teaching practices and create a more engaging learning atmosphere. The study has shown that when learners are satisfied with their learning experiences and enjoy using the chatbots, their engagement and CU are enhanced; hence, teachers should consider using these teaching strategies.

Author contributions: SJ & MM: writing – original draft, writing – review & editing; **SJ:** data curation, formal analysis; **MM:** conceptualisation. Both authors have agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study was approved by the Research Ethics Committee at the University of Venda No FHSSE/23/PCEM/03/3008. Written informed consents were obtained from the participants.

AI statement: The authors stated that an AI-based grammar checker, Grammarly, was used for proofreading. No content generation was performed by AI.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Abuhassna, H., Al-Rahmi, W. M., Yahya, N., Zakaria, M. A. Z. M., Kosnin, A. B. M., & Darwish, M. (2020). Development of a new model on utilizing online learning platforms to improve students' academic achievements and satisfaction. *International Journal of Educational Technology in Higher Education, 17*, Article 38. https://doi.org/10.1186/s41239-020-00216-z

Al Faraby, S., Adiwijaya, A., & Romadhony, A. (2023). Review on neural question generation for education purposes. *International Journal of Artificial Intelligence in Education, 34*, 1008-1045. https://doi.org/10.1007/s40593-023-00374-x

Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Research in Science Education, 54*(5), 977-997. https://doi.org/10.1007/s11165-024-10176-3

Al-Momani, M. M., & Pilli, O. (2021). The impact of integrated mobile system on students' academic achievement, interaction and satisfaction. *TEM Journal, 10*(3), Article 1395. https://doi.org/10.18421/tem103-50

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52-62. https://doi.org/10.61969/jai.1337500

Bandura, A. (2012). Cultivate self-efficacy for personal and organizational effectiveness. In E. A. Lacke (Ed.), *Handbook of principles of organizational behavior* (pp. 179-200). Wiley. https://doi.org/10.1002/9781119206422.ch10

Bandura, A., Freeman, W. H., & Lightsey, R. (1999). Self-efficacy: The exercise of control. *Journal of Cognitive Psychotherapy, 13*(2), 158-166. https://doi.org/10.1891/0889-8391.13.2.158

Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C. D. (2018). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology, 53*, 87-105. https://doi.org/10.1016/j.cedpsych.2018.01.002

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), Article 588. https://doi.org/10.1037//0033-2909.88.3.588

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, 32*(7), 3240-3256. https://doi.org/10.1080/10494820.2023.2172044

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32*(3), 444-452. https://doi.org/10.1007/s10956-023-10039-y

Department of Basic Education. (2011). Curriculum and assessment policy statement grades 10-12 physical sciences. *Department of Basic Education, South Africa*. https://www.education.gov.za/LinkClick.aspx?fileticket=uVcOcx728Y8%3D&tabid=466

Department of Basic Education. (2018). Grade 11 physical sciences: Chemistry (P2). *Department of Basic Education, South Africa*. https://www.education.gov.za/Curriculum/NationalSeniorCertificate(NSC)Examinations/Grade11Exams.aspx

dos Santos, R. P. (2023). Enhancing chemistry learning with ChatGPT, Bing Chat, Bard, and Claude as agents-to-think-with: A comparative case study. *SSRN*. https://doi.org/10.2139/ssrn.4629567

Du, B. (2023). Research on the factors influencing the learner satisfaction of MOOCs. *Education and Information Technologies, 28*(2), 1935-1955. https://doi.org/10.1007/s10639-022-11269-0

Fornell, C., & Bookstein, F. L. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research, 19*(4), 440-452. https://doi.org/10.1177/002224378201900406

Gill, S. S., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., & Parlikad, A. K. (2024). Transformative effects of ChatGPT on modern education: Emerging era of AI chatbots. *Internet of Things and Cyber-Physical Systems, 4*, 19-23. https://doi.org/10.1016/j.iotcps.2023.06.002

Hair Jr, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. (2014). *Multivariate data analysis*. Pearson.

Hair Jr, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2-24. https://doi.org/10.1108/ebr-11-2018-0203

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), Article 424. https://doi.org/10.1037//1082-989x.3.4.424

Hua, M., Wang, L., & Li, J. (2024). The impact of self-directed learning experience and course experience on learning satisfaction of university students in blended learning environments: The mediating role of deep and surface learning approach. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1278827

Iyamuremye, A., & Ndihokubwayo, K. (2024). Exploring secondary school students' interest and mastery of atomic structure and chemical bonding through ChatGPT. *Educational Journal of Artificial Intelligence and Machine Learning, 1*(1), 1-13. https://doi.org/10.58197/prbl/9hk37296

Jacobsen, L. J., & Weber, K. E. (2023). The promises and pitfalls of ChatGPT as a feedback provider in higher education: An exploratory study of prompt engineering and the quality of AI-driven feedback. *OSF*. https://doi.org/10.31219/osf.io/cr257

Jere, S. (2025). Evaluating artificial intelligence large language models' performances in a South African high school chemistry exam. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(2), Article em2582. https://doi.org/10.29333/ejmste/15932

Jere, S., Bessong, R., Mpeta, M., & Litshani, N. F. (2024). Exploring pre-service teachers' perceptions of ChatGPT integration into physical sciences teaching: A case study at a rural South African university. *International Journal of Learning, Teaching and Educational Research, 23*(11), 464-486. https://doi.org/10.26803/ijlter.23.11.24

Jia, X.-H., & Tu, J.-C. (2024). Towards a new conceptual model of AI-enhanced learning for college students: The roles of artificial intelligence capabilities, general self-efficacy, learning motivation, and critical thinking awareness. *Systems, 12*(3), Article 74. https://doi.org/10.3390/systems12030074

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2020). Damage caused by women's lower self-efficacy on physics learning. *Physical Review Physics Education Research, 16*, Article 010118. https://doi.org/10.1103/physrevphyseducres.16.010118

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*(3), 157-163. https://doi.org/10.1177/0963721412443552

Karpicke, J. D., & Grimaldi, P. J. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review, 24*(3), 401-418. https://doi.org/10.1007/s10648-012-9202-2

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kıyak, Y. S. (2023). A ChatGPT prompt for writing case-based multiple-choice questions. *Revista Española de Educación Médica, 4*(3), 98-103. https://doi.org/10.6018/edumed.587451

Kolil, V. K., Muthupalani, S., & Achuthan, K. (2020). Virtual experimental platforms in chemistry laboratory education and its impact on experimental self-efficacy. *International Journal of Educational Technology in Higher Education, 17*, Article 30. https://doi.org/10.1186/s41239-020-00204-3

Küchemann, S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). Can ChatGPT support prospective teachers in physics task development? *Physical Review Physics Education Research, 19*, Article 020128. https://doi.org/10.1103/physrevphyseducres.19.020128

Lane, S., Hoang, J. G., Leighton, J. P., & Rissanen, A. (2021). Engagement and satisfaction: Mixed-method analysis of blended learning in the sciences. *Canadian Journal of Science, Mathematics and Technology Education, 21*(1), 100-122. https://doi.org/10.1007/s42330-021-00139-5

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), Article 16. https://doi.org/10.1186/s41239-024-00447-4

Lee, J., An, T., Chu, H.-E., Hong, H.-G., & Martin, S. N. (2023). Improving science conceptual understanding and attitudes in elementary science classes through the development and application of a rule-based AI Chatbot. *Asia-Pacific Science Education, 9*(2), 365-412. https://doi.org/10.1163/23641177-bja10070

Li, Z., & Li, J. (2022). Using the flipped classroom to promote learner engagement for the sustainable development of language skills: A mixed-methods study. *Sustainability, 14*(10), Article 5983. https://doi.org/10.3390/su14105983

Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching, 6*(1), 64-74. https://doi.org/10.37074/jalt.2023.6.1.32

Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *The Journal of Academic Librarianship, 49*(4), Article 102720. https://doi.org/10.1016/j.acalib.2023.102720

Long, H. B. (1985). Contradictory expectations? Achievement and satisfaction in adult learning. *The Journal of Continuing Higher Education, 33*(3), 10-12. https://doi.org/10.1080/07377366.1985.10401035

Marsh, H. W., & Balla, J. (1994). Goodness of fit in confirmatory factor analysis: The effects of sample size and model parsimony. *Quality and Quantity, 28*(2), 185-217. https://doi.org/10.1007/bf01102761

Nguyen, G., Nguyen, N., & Giang, N. (2022). Situation and proposals for implementing artificial intelligence-based instructional technology in Vietnamese secondary schools. *International Journal of Emerging Technologies in Learning, 17*(18), 53-75. https://doi.org/10.3991/ijet.v17i18.31503

Polakova, P., & Klimova, B. (2024). Implementation of AI-driven technology into education–a pilot study on the use of chatbots in foreign language learning. *Cogent Education, 11*(1), Article 2355385. https://doi.org/10.1080/2331186x.2024.2355385

Qawaqneh, H., Ahmad, F. B., & Alawamreh, A. R. (2023). The impact of artificial intelligence-based virtual laboratories on developing students' motivation towards learning mathematics. *International Journal of Emerging Technologies in Learning, 18*(14), Article 105. https://doi.org/10.3991/ijet.v18i14.39873

Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies, 26*(3), 2623-2656. https://doi.org/10.1007/s10639-020-10375-1

Rizvi, S. (2023). Revolutionizing student engagement: Artificial intelligence's impact on specialized learning motivation. *International Journal of Advanced Engineering Research and Science, 10*(9). https://doi.org/10.22161/ijaers.109.4

Sarstedt, M., Ringle, C. M., & Hair Jr, J. F. (2021). Partial least squares structural equation modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Handbook of market research* (pp. 587-632). Springer. https://doi.org/10.1007/978-3-319-57413-4_15

Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly, 23*(1), 7-25. https://doi.org/10.1080/10573560600837578

Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and Psychological Measurement, 80*(3), 421-445. https://doi.org/10.1177/0013164419885164

Spector, P. E. (2019). Do not cross me: Optimizing the use of cross-sectional designs. *Journal of Business and Psychology, 34*(2), 125-137. https://doi.org/10.1007/s10869-018-09613-8

Van der Stede, W. A. (2014). A manipulationist view of causality in cross-sectional survey research. *Accounting, Organizations and Society, 39*(7), 567-574. https://doi.org/10.1016/j.aos.2013.12.001

Wang, M.-T., Fredricks, J. A., Ye, F., Hofkens, T. L., & Linn, J. S. (2016). The math and science engagement scales: Scale development, validation, and psychometric properties. *Learning and Instruction, 43*, 16-26. https://doi.org/10.1016/j.learninstruc.2016.01.008

Wang, Y.-Y., & Chuang, Y.-W. (2024). Artificial intelligence self-efficacy: Scale development and validation. *Education and Information Technologies, 29*(4), 4785-4808. https://doi.org/10.1007/s10639-023-12015-w

Wu, T.-T., Lee, H.-Y., Li, P.-H., Huang, C.-N., & Huang, Y.-M. (2024). 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

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education, 9*, Article 59. https://doi.org/10.1186/s40594-022-00377-5

Yu, Z., Gao, M., & Wang, L. (2021). The effect of educational games on learning outcomes, student motivation, engagement and satisfaction. *Journal of Educational Computing Research, 59*(3), 522-546. https://doi.org/10.1177/0735633120969214

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166-183. https://doi.org/10.3102/0002831207312909