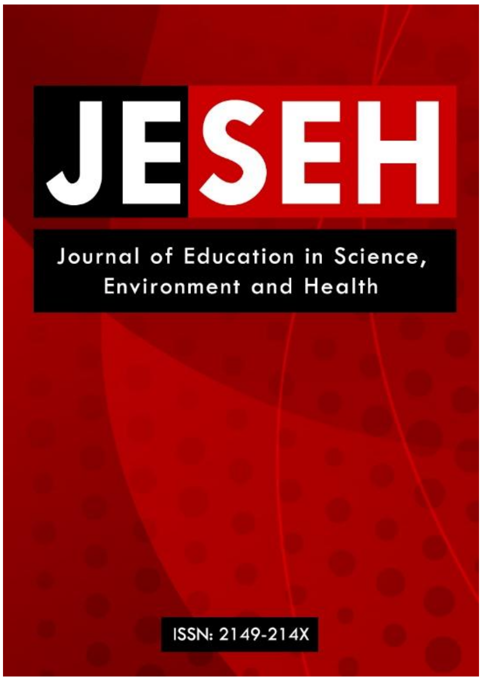

ISSN: 2149-214X

Journal of Education in Science, Environment and Health

www.jeseh.net

Teaching and Learning Chemistry for the 21st Century Skills Through Artificial Intelligence - A Narrative Review

Tebogo E. Nkanyani
University of Pretoria

To cite this article:

Nkanyani, T.E. (2026). Teaching and learning chemistry for the 21st century skills through artificial intelligence: A narrative review. *Journal of Education in Science, Environment and Health (JESEH), 12*(2), 198-208. https://doi.org/ 10.55549/jeseh.898

Journal of Education in Science, Environment and Health
Volume 12, Issue 2, 2026 https://doi.org/10.55549/jeseh.898

Teaching and Learning Chemistry for the 21st Century Skills Through Artificial Intelligence - A Narrative Review

Tebogo E. Nkanyani

Article Info

Article History

Published:
01 April 2026

Received:
06 February 2026

Accepted:
03 March 2026

Keywords

Generative artificial intelligence
Critical thinking
Misconceptions

Abstract

This review provides an overview of how GAI assists in chemistry learning and teaching. Aspects such as 21st century skills and the tutoring ability of GAI were examined from the learning point of view, while also covering pedagogical aspects such as Technological Pedagogical Content Knowledge (TPACK), visualization and representation levels of Chemistry, textual aspects, and problem-based learning (PBL). The GAI seemed to elicit students' 21st century skills and, to a lesser extent, display a tutoring ability, depending on whether it was a free or paid version. Furthermore, the GAI requires objectivity and sufficient TPACK owing to its hallucinations and inconsistencies in definitions, diagram interpretation, and image generation, among others. Furthermore, the chatbots seemed to struggle with representation levels, generating responses with macro-level definitions for concepts requiring definitions at the sub-micro level. These inconsistencies and hallucinations, without meticulous verification, can lead to misconceptions. Moreover, social inequalities may negatively impact meaningful learning, as paid chatbots perform better than free versions in generating and interpreting chemistry images, among other abilities.

Introduction

Artificial Intelligence (AI) has significantly impacted the education fraternity, with tools such as Chat Generative Pre-Trained Transformer (ChatGPT), notably the frequently used chatbot. Introduced late in 2022 (Haleem et al. 2022), ChatGPT has made its mark in teaching and learning. It displays many abilities, such as assisting teachers with lesson planning (Sykes, 2023), producing classroom materials (Valeri, 2025), and designing assessments (Wandi et al., 2025). However, this ability is not limited to ChatGPT and is demonstrated by other AI software. For example, Google Bard, among others, has the ability to assist students in enhancing their writing (Ling Jen & Salam, 2024). Nonetheless, it is important that students learn and improve their writing throughout the process. The advantage of AI for the learning process is that it can produce excellently written text that looks overboard and error-free. AI also provides platforms for engagement between the user and the AI through prompting engineering (Fenta, 2025). Furthermore, AI chatbots assist in reducing teachers' workload through creation, grading, and informing students (Labadze et al., 2023). Nevertheless, ethical issues must be considered. For instance, students may use AI to write assignments without actively engaging in the learning process. A story by the Independent Online (IOL) newspaper from South Africa highlights the concern demonstrated by academics, who, apart from acknowledging the significance of AI in education, continue to note the misuse of AI chatbots by university students through the 'copy and paste' technique without engaging with the chatbot in a way that will make learning meaningful (Buthelezi, 2025). Students may be taking that approach as a shortcut or due to their inability to deal with assignment loads and deadlines, consequently resorting to that unethical approach. Some take this approach because of the challenging nature of science and the complexity of some chemistry topics. Generative AI (GAI) also has other shortcomings, such as gender bias and hallucinations (Feldman-Maggor et al., 2025), consequently providing irrelevant and incorrect responses (Sedagat, 2025; Elmas et al., 2024). Therefore, users must be well conversant with the content they are interacting with on AI chatbots to avoid collecting incorrect or incomplete information.

Current Review

Understanding the use of generative artificial intelligence (GAI) in chemistry education is important as it can inform teachers and all role players about effective innovative teaching practices. Previous literature has delved into this area, but not to the extent of this review. For example, Pabuçcu-Akış (2024) conducted a bibliometric analysis of innovative technologies in organic chemistry. The author found AI as a recent tool that was implemented in organic chemistry at a rate of 16,7 %. Nonetheless, there was no clarity on how AI was used in the reviewed studies. Similarly, Iyamuremye et al. (2024) explored the use of AI and machine learning (ML) in chemistry education. The study found opportunities for the use of AI and ML, such as individualized instruction, support from teachers, and availability of educational resources, while it also noted challenges such as reliance on available data, bias in the available models, and privacy and security issues regarding data use. It was also highlighted that the platform was used in instances of curricular development, lesson preparation, and student assessment. However, there was no clarity on how AI was used and on which components it was helpful. This review takes things further by exploring the manner in which GAI was implemented in Chemistry Education, with a focus on the learning process and pedagogical practices. The intention was to first understand how GAI influenced the learning aspects of chemistry from the students` perspective. The focus was on students' 21st century skills, and the tutoring abilities of GAI, as described in the next section. The second aspect of the review was how teachers used GAI for teaching. Instead of looking at how GAI assisted the chemistry teacher in aspects of lesson planning and assessment, this review went deeper into other critical aspects such as teachers' technological pedagogical content knowledge (TPACK), chemistry visualization, representation levels of chemistry, textual aspects of chemistry, and problem-based learning (PBL).

Learning Chemistry Through GAI

21st Century Skills

The 21st century skills are essential for the growth of any country's economy as they are consistently used in industries and other sectors. These include critical thinking, problem solving, information analysis, reasoning, and inquiry (Binkley et al., 2012). Haryani et al. (2024) expand on this by highlighting the 4Cs: "creativity, critical thinking, problem solving, communication, and collaboration" (p. 106). These skills, which are critical for the chemical industry, are emphasized in most school curricula in different countries. For example, the South African Physical Sciences Curriculum And Policy Statement (CAPS) encourages as one of its aims, "identify and solve problems and make decisions using critical and creative thinking" (DBE, 2011, p.5). On the same page (page 5), CAPS outlines aim that align with 21st century skills. In science education, there is plenty of literature on 21st century skills. For example, studies by Haryani et al. (2024), Amandi (2023), and Kennedy (2022). Nevertheless, my interest lies in the use of AI to elicit 21st century skills in chemistry education.

Critical Thinking Skills

Authors such as Cooper (2023) and Lawasi et al. (2024) have highlighted critical thinking as an important skill that students should acquire when learning science through GAI. In this context, students would ask specific questions and consequently take a critical posture in the responses generated by the chatbot (Lawasi et al., 2024). Guo and Lee (2023) explored how the use of AI chatbots, such as ChatGPT, influences students' critical thinking skills. In essence, students were not only expected to use ChatGPT for assistance with schoolwork, but also to interact with the chatbot to consequently "critique, evaluate, analyze, and draw logical conclusions" (p.4844). This is crucial for the chemistry learning process, as students would not just take information as it comes, but make judgments as they interact with the chatbot, choosing what is important and sensible. Particularly with the inaccuracies that chatbots such as ChatGPT bring through their responses (Sedagat, 2025; Elmas et al., 2024). The study (Guo & Lee, 2023) noted positive views from students on their ability to pose challenging questions, assess information, summarize it, and solve difficult content. The demonstration of critical thinking skills by students would assist them in developing objectivity rather than subjectivity when they have information at their disposal. In another study by Tassotti (2024), students noticed the incorrectness of responses from ChatGPT at a higher chemistry level. This has positive effects on the pedagogical abilities of these future teachers, as the presence of critical thinking would enable them to make better judgements on what to use and not to use in their teaching and assessment.

Problem Solving

The other 21st century skill that was explored was problem-solving. Clark (2023) used the problem-solving model initially proposed by Taasoob and Glynn (2009) to analyze ChatGPT's closed-ended responses on chemistry topics. The model would have expected the chatbot to follow the sequence: problem conceptualization, problem strategy, and then finalize with problem solution and assessment. However, the chatbot was not effective for this task. In fact, ChatGPT was only effective in problem conceptualization of both numeric and non-numeric words. Nonetheless, the chatbot did not perform well in coming up with a problem strategy, especially for numeric questions on titration. The chatbot erred by first calculating the number of moles using the formula moles = concentration × volume, but not first converting the volume from milliliters to liters, resulting in the calculation of millimoles instead of moles (Clark, 2023). This ties in with the critical thinking skills indicated above, as students will be required to note such inaccuracies after proper assessment of responses. Further, in the example given above (Tassotti, 2024), students were able to adjust their prompts through the 5S strategy: "set the scene, simplify your language, share feedback, structure the output, be specific" (p. 2467). Through this adjustment, they were able to see improvements in the ChatGPT response and were satisfied. This notable improvement in the prompts signals a problem-solving skill, which may be key in their chemistry teaching practice. In contrast, ChatGPT 4.0 seemed to be doing better by providing an organized dissection of the interpretation of the free energy diagrams, with positive implications for critical thinking and problem-solving skills (Alasadi & Baiz, 2024). However, ChatGPT 4.0 is not a free version and requires a subscription.

Information Analysis and Reasoning

Information analysis is a critical aspect of the learning process. With this skill, students are able to critically judge what is in front of them, be it boiling points of different substances, bond energies and enthalpy values, or any other chemistry information. But how do chemistry students interact with the information provided by GAI? Ruff et al. (2024) provided a rapport of the influence of ChatGPT in student learning. Biochemistry Lab and inorganic students were able to note how ChatGPT assisted in gramma revisions. However, they also noted its inability to generate formulaic structures, and consequently suggested how ChatGPT's generative ability could be improved. A crucial aspect that ties in with their critical thinking and reasoning skills. Students in the study by Tassoti (2024) used one of the components of the five S prompting strategy, 'share feedback, ' to interact with ChatGPT to attempt to modify a certain aspect of the answer generated. They used phrases such as "please include the same facts for sulfur" (p. 2479). This type of interaction, for example, is indicative of students' ability to analyze and make judgements on information given and, more importantly, make further refinements. Other students demonstrated this by comparing the responses of ChatGPT with other sources. For example, in the same study by Tassoti (2024), a student used the following as a follow-up prompt: "I have a source saying there are 6 oxidation levels, but you say there are 4. Which statement is legit?" (p. 2479). This is indicative of a student who is willing to interact with GAI instead of copying and pasting information as it comes.

Inquiry

As a hands-on subject, science is embedded in inquiry. If used effectively, inquiry can elicit students' critical thinking skills, creativity, and innovation, among other skills. In chemistry education, students may be required to undergo different stages of inquiry, which may result in a meaningful learning experience. Pedaste et al. (2015, p.54) identify "orientation, conceptualization, investigation, conclusions, and discussion" as the general phases of inquiry. However, the phases may differ based on students' levels. Of interest to this review is whether students can use GAI for inquiry-based learning in chemistry education. Kim (2025) undertook a study wherein college students were required to produce cyclohexene from cyclohexanol. The students employed AI chatbots, such as ChatGPT, Gemini, and Microsoft Copilot, to find greener alternatives to normal reagents. The activity yielded students' digital literacy while observing sustainability and science inquiry. Students were exposed to prompt engineering, where their interaction with the chatbot improved their engagement while shaping their critical thinking skills, as demonstrated above. Although the study primarily focused on abstract chemistry, the findings have positive implications for chemistry education.

Tutoring Ability of ChatGPT Free Version

Just like any other resource, students would engage with ChatGPT in the same way they interact with textbooks, websites, YouTube videos, and other available resources. But what is crucial for this study is how the chatbot assumes a tutoring role. Leon and Vidhani (2023) explored this arena by focusing on the patterns of prompting, response coherence, and trustworthiness of chatbots. ChatGPT provided different answers for the same multiple-choice question on the number of potassium core electrons when asked repeatedly several times. This highlights its lack of reliability in answering chemistry multiple-choice questions, which can have a detrimental impact on the learning process in cases where students are dependent on it for correct answers (Leon & Vidhani, 2023). Moreover, Tassoti (2024) highlighted how ChatGPT hallucinates and struggles to provide correct responses to high-order questions. This shortcoming may have detrimental effects on advanced students who always want to challenge themselves with more complex content.

Furthermore, in addition to being biased, hallucinations may lead students, especially those who are overly reliant on resources, to take in incorrect information (Feldman-Maggor et al., 2024), attracting misconceptions in the process. ChatGPT 4.0 was also found to demonstrate tutoring abilities by answering questions that require graphical analysis. For example, in the study by Alasadi and Baiz (2024), the chatbot was asked to analyze the free energy diagram, calculate the thermodynamic favorability of the reactant or product using the free energy change for both plots provided, and evaluate how the amount of activation energy affects the rate of reaction. Consequently, the chatbot was able to correctly distinguish between the two plots, indicating the first plot as that of that of a negative $\triangle G$ which is in contrast to the second plot with positive $\triangle G$. Moreover, the chatbot was able to indicate the limitations of the plots, for example: "these diagrams do not provide quantitative kinetic data but allow for the comparative understanding of the relative activation energies (E_a) and the potential rate of reaction," (p. 2718). This sense of transparency has positive consequences for the learning process, as students would be able to understand what each image represents and what it cannot. Nonetheless, the authors felt it could do better in highlighting possible misconceptions regarding the interpretation of plots and determination of results, which could be useful for first-time learning (Alasadi & Baiz, 2024). Despite these shortcomings, ChatGPT4.0 still demonstrated tutoring abilities similar to those of "personal tutors" through its ability to provide quick and sufficient clarity of different via a number of follow-up inquiries (Alasadi & Baiz, 2024, p. 2719). Its tutoring capabilities in analyzing challenging chemistry diagrams are indicative of knowledge advancement in chemistry concepts, which may assist in the student's conceptual understanding through engagement with the chatbot on scientific information and diagrams (Alasadi & Baiz, 2024).

The tutoring capability of ChatGPT 4.0 was again tested in organic chemistry diagrams of different qualities, where the first one was assisted through hand annotations, while the second and third diagrams were hand drawn, with the third adjusted to show double bonds. The chatbot was asked to solve the synthesis problem, provide details of how to approach similar problems, provide reagents required for each step, and make judgements on the order of reagents versus the expected outcomes (Alasadi & Baiz, 2024). ChatGPT 4.0 was able to provide clear step-by-step instructions for the synthesis of meta-bromo nitrobenzene and provided clear justification for the positioning of bromination, showing its tutoring efficiency by allowing students to mentally sequence organic reaction transformation, yielding a deeper interaction with the content (Alasadi & Baiz, 2024). Moreover, ChatGPT 4.0 was able to maintain its objective of explaining concepts despite a change in the quality or orientation of the image. However, its shortcoming was that it misinterpreted the images with "reduced resolution and inversion," incorrectly labelling them as bromo-nitrobenzene, while its correct identification is a meta-isomer (p. 2720). Furthermore, the chatbot's interpretation of the two drawn diagrams mentioned above was reported to provide different and incorrect solutions to the synthesis problem (Alasadi & Baiz, 2024). This would require a high level of objectivity from students, rather than subjectively relying on ChatGPT responses.

Teaching Chemistry Through GAI

It is also crucial to understand how GAI assists in chemistry teaching. However, teaching is complex and can be viewed from different perspectives. This review focuses on TPACK, visualization of chemistry, representation level of chemistry, textual aspects of chemistry, and problem-based learning (PBL).

TPACK

Feldman-Maggor et al. (2024) explored the type of knowledge required for teachers to efficiently use GAI. TPACK was used as a lens because of its nature of assessing teacher knowledge and pedagogies from a technological point of view. Even though authors such as Lorenzo and Romeike (2023) suggested an extension of TPACK with an AI component to yield DTPACK, disregarding the invalidity of TPACK alone to measure the use of GAI, the authors Feldman-Maggor still felt its relevance in their study. Nonetheless, they echoed the sentiments of Lorenzo and Romeike. Their intention was to understand how teachers CK and PCK influence their prompting ability and how they would ultimately assess and judge the quality of the prompts generated. This is a crucial aspect of teaching and has strong implications for science lesson planning, as teachers are expected not only to acquire different sources but also to sort what is important from what is not relevant to their lesson. For example, they used a teacher's dialogue with ChatGPT to assess whether the teacher would be able to use PCK to evaluate the response generated by the chatbot and ultimately follow up with content-specific prompts. The dialogue was about the teacher asking for the difference between ionic and molecular materials and ultimately seeking guidance on relevant teaching strategies. The teacher sought questions that would assist in eliminating students' misconceptions about NaOH and CH_3CH_2OH, regarding the role of "OH" as an ion and -OH as a functional group. Although the chatbot did not go where the teacher wanted it to, the teacher modified and improved the prompts using their own PCK.

The chatbot provided a satisfactory answer, although it was still insufficient. The chatbot also hallucinated by providing incorrect references that could not be identified after the teacher's verification. Therefore, the teacher demonstrated sufficient PCK and CK through dialogue with ChatGPT prompts. The implication is that the 'T' part of TPACK aligns to the teachers' PCK. This underscores the importance of teachers' PCK when seeking clarity from ChatGPT in their lesson preparation, which would allow them to be objective and critical throughout the process. The absence of such a framework may lead to teaching lessons that are fully characterized by misconceptions.

Visualization of Chemistry

Visualization is important in chemistry because it has a bearing on students' understanding of chemistry through its different levels of representation: macro, sub-micro, and symbolic (Talanquer, 2011). I am interested in the image production capabilities of generative AI in chemistry education. A chatbot, such as ChatGPT 4.0, was tested by Alasadi and Baiz (2024) to assess its ability to engage with and interpret visual chemistry representations. In this context, the study explored the tutoring abilities of the chatbot through engagement with students, the chatbot's ability to provide understanding of complex figures, analysis of chemical structure, and ability to read tables and graphs, among others (Alasadi & Baiz, 2024). These aspects are key to the conceptual understanding of chemistry, which is embedded in visual representations. However, even though ChatGPT was able to interpret basic chemical structures, it struggled with complex diagram analysis when images were inverted or diagrams were of poor quality. Nonetheless, the paid version, ChatGPT 4.0, was efficient in reading hand-drawn tables and graphs, students' handwritten calculations, and spectral analysis with favorable effects on organic chemistry – reading molecular formulas and molecular structures where it showed variation in isomers and identifying relevant structures (Alasadi & Baiz, 2024). Nonetheless, the implications of these abilities for teaching chemistry still require further exploration. Nascimento Júnior et al. (2024) went a step further by exploring the text and image capabilities of seven GAI bots such as ChatGPT 3.5; ChatGPT 4.0, Google Bard, Bing Chat, Adobe Firefly, Leonardo AI, DALLE-E in chemistry education. The focus was on how these GAIs generate textual and image content of chemistry. In this subsection, I focus on the implications for the visual aspects of teaching chemistry. In this context, the GAI chatbots were exposed to text through prompting and were expected to generate images; they also used the same prompts to recognize images from organic chemistry reaction mechanisms, resonance, and energy diagrams. The image creation aspects were chemical bonds, Lewis structures, and atomic models. However, the authors noted that chatbots such as Copilot/BingChat and Google Bard/Gemini were unable to respond effectively to the prompts compared with ChatGPT 4.0, which appeared to be responding well to text despite some notable errors. ChatGPT 4.0 was able to generate correct chemical responses, accurately recognizing molecules and forms of arrows compared to Copilot/BingChat and Google Bard/Gemini. Nascimento Júnior et al. (2024) further indicated pointed that Google Bard, unlike its counterparts, failed to identify alkyl bromide as the reaction substrate. The GAI chatbots also seemed to struggle with generating covalent bond representations (Nascimento Júnior et al., 2024). The representation had a narrow nucleus closely attached to electrons that seemed to share orbits, which could create misconceptions. The GAI also uses the concept of electron transfer to represent ionic bonds rather than demonstrating ionic bonding itself. Nonetheless, the GAI performed well when the user changed from DALL-e to Python code while prompting for Lewis structures. Furthermore, Akaygun and Kilic (2025) conducted a study in which chemistry PSTs were expected to use chatbots to create visualizations that could be used in their lessons. PSTs were able to apply their prompting abilities influenced by their PCK, which they adapted to achieve desirable outcomes.

Representation Levels of Chemistry

Chemistry is represented as either or a combination of symbolic, macroscopic, and/or sub-microscopic levels. However, students still find it difficult to integrate the levels in chemistry learning (Murni et al., 2022). They confuse one level as the other level and vice versa (Nascimento Junior et al., 2024). The GAI prompts responses seem to be of less help because of their definitions of covalent, metallic, and ionic bonds. They would provide a definition for one type that focuses on one type of representation, while the definition for the other type of bond will focus on other representation types. This may create more confusion for students when one teaches them with unverified content from chatbots, where they may end up seeing the same representations in different definitions, while that is not supposed to be the case, building a myriad of misconceptions and misunderstandings in the process. Furthermore, Nascimento Júnior et al. (2024) found that GAI responses failed to show an understanding of different representations on more than one occasion. For example, when illustrating metallic bonds, the GAI used macroscopic characteristics such as color and texture instead of representations from a sub-microscopic level.

Textual Aspects of Chemistry

The GAI seems to perform well in generating text from chemistry content. In their study, Nascimento Junior et al. (2024) applied prompting engineering to seek a definition for covalent bonding. Even though the chatbot provided a well-written definition, it contained some inaccuracies that may have led to students' misconceptions. Phrases such as "covalent bonds typically form between non-metal atoms, as these elements tend to gain, loose or share electrons…." (p. 3770) may lead to students mistaking ionic and covalent bonds (Nascimento Junior et al., 2024). Furthermore, all the chatbots under study seemed to "personify" atoms in their definition of covalent bonds by using words such as "atom seek" (Nascimento Junior et al., 2024, p. 3770). The language produced by the chatbots may, in such cases, lose its scientific nature, as chemistry involves non-living particles. Moreover, the chatbots such as Google Bard, Bing chat, ChatGPT 3.5, and ChatGPT 4.0 gave responses that were common with students' misconceptions on chemical bonding, irrespective of the level of complexity of the prompts generated – from beginner, intermediate, then advanced prompts (Nascimento Junior et al., 2024). Therefore, science teachers must be meticulous when using chatbots to prepare teaching content. Teachers would have to use different sources to cross-compare the responses from the chatbots.

Problem Based Learning (PBL)

As an active learning strategy, PBL is a highly encouraged teaching strategy in science. This is because it complements the inquiry nature of science. Furthermore, this aligns well with the aims of the science curriculum. For example, in the South African curriculum and policy statement (CAPS) for physical sciences, skills such as problem-solving and critical thinking are over-emphasized under the assessment taxonomy' and 'aims and principles' sections. Equally, it is prudent to understand how the use of GAI assists in the implementation of those skills in the science classroom. The literature shows that through the use of AI simulations, students are more interactive, and learning objectives are effectively attained (Ramos & Condotta, 2024). In this context, students were able to simulate the drying curves and particle size distributions using ChatGPT. Even though the study involved chemical engineering students, there are key highlights that the study makes which has implications for pedagogical aspects of chemistry education. For example, in one of the ChatGPT prompts, the participants indicated their intention to design a PBL task and asked the chatbot to suggest a topic (p. 3249). However, this was not the end of the story. The participants asked the chatbot to suggest extra 7 topics in order to accommodate all eight groups. In summary, not only did ChatGPT assist practitioners in the conceptualization part of the PBL task, but also shared pedagogical responsibilities, such as producing first drafts, generating rubrics, inquiry stages, preliminary simulation settings, and team aspects (Ramos & Condotta, 2024). The implication is that chemistry teachers would be able to use ChatGPT in the design and assessment of PBL tasks, whether in the acid and bases topic, organic compounds, or even separation of substances topics at the senior high school level. The ability of ChatGPT in the conceptualization point will come in handy in helping the teacher select and allocate different subtopics for each group. However, teachers would still have to apply their PCK and skills, using ChatGPT as a co-teacher rather than a complete replacement of their involvement. This step would be useful during the prompting stage.

Discussion

Learning Chemistry Through GAI

Learning is one of the most critical components of education, and scholars have always defined and redefined how different teaching strategies and platforms influence the learning process. This review highlighted how GAI influences 21st century skills in the learning of chemistry. Important skills, such as critical thinking skills, have been highlighted in this study, where students, through their expressions, are able to assess the information provided by the chatbots, challenge it with questions, summarize the information, and use it to solve complex subject matter (Guo & Lee, 2023). This skill is also demonstrated when students are able to notice the inaccuracy of the information a chatbot like ChatGPT would bring in the higher chemistry level (Tassotti, 2024). Inaccuracies, if not noted and addressed in the learning process, may lead to misconceptions, which is ineffective for the learning process. Therefore, critical thinking skills are key in the use of GAI for learning chemistry, as they would allow students to be objective in the information they receive. However, objectivity also requires students to have information analysis and reasoning skills. Similar to critical thinking, students are expected to be critical of the information at their disposal, which in this case would be AI responses. Students noted that even though ChatGPT assisted in their revision of grammar, it could not produce formulaic structures (Ruff et al., 2024). Formulae are important in chemistry learning as they play a role in understanding reactions and mechanisms. It is also critical in the representation of chemistry, which is embedded in the particle and molecular nature. Through their information analysis and reasoning, students were able to suggest how the chatbot's generative ability could be improved, which has positive pedagogical implications for student teachers who would use that skill during their teaching practices. Chemistry students can demonstrate their information analysis and reasoning skills through the 5S prompting strategies (Tassoti, 2024). In the process, the students were able to modify their prompts through their interaction with ChatGPT while verifying the responses from ChatGPT through other sources. Interestingly, they used the information they received from other sources to create a dialogue with the chatbot, which is crucial for meaningful learning that encourages student-centeredness. Teachers should teach chemistry in a way that encourages the development of these crucial critical thinking, information analysis, and reasoning skills, which have an effect on learning on platforms other than the classroom. More importantly, environments which would require students to be self-directed. Moreover, chemistry learning should also take into cognizance that science is a hands-on subject that is embedded in inquiry and problem-solving part the key skills. Students can use GAI for the conceptualization of the inquiry, where they use AI chatbots to suggest alternative reagents (Kim. 2025). These have positive implications for primary and secondary school students who may use GAI to conceptualize their inquiry projects, where they are required to, for example, find recyclable material and do water quality projects, among others. However, they may require guidance from their teachers, depending on the classroom level. Nonetheless, students can integrate their problem-solving strategies during inquiry through GAI (Clark, 2023). However, a chatbot like ChatGPT has some limitations in that it is only able to conceptualize the problem but is unable to generate a problem strategy for topics such as titration when it involves numeric words (Clark, 2023). Therefore, students need to be sharp in their problem solving when using GAI to conceptualize chemistry inquiry. Learning through the GAI also has implications that may require chatbots to assume the tutor role, especially when students learn in their own space without the presence of their teachers. The chatbot, such as ChatGPT, was found to be inconsistent in its responses when asked the same question a couple of times (Leon & Vidhani, 2023) and gave incorrect answers for high-order questions, hallucinating in the process (Tassoti, 2024). This shortcoming negatively impacts chemistry learning, especially for students who challenge themselves with higher-order questions during the learning process. Moreover, hallucinations may add to the misconceptions that students already have regarding the topics they are learning. Nonetheless, ChatGPT 4.0 seems to perform better than its predecessor. For example, it can analyze graphical information from the free energy diagram, calculate the thermodynamic favorability, and measure the effect of activation energy on the rate of reaction (Alasadi & Baiz, 2024). Nonetheless, it could not highlight possible misconceptions in the data. The chatbot demonstrated its tutoring abilities through organic chemistry diagrams of different annotation qualities and asked to provide reagents, order, and outcomes, which it did not disappoint (Alasadi & Baiz, 2024). Nonetheless, it struggled with low-quality and low-resolution diagrams. These shortcomings would require students to have a combination of the 21st century skills mentioned above, tied with objectivity. Moreover, even though ChatGPT 4.0 provides a better tutoring ability than ChatGPT 3.5, it is a paid chatbot that requires users to subscribe, limiting access. In particular, students from more disadvantaged backgrounds are found.

Teaching Chemistry Through GAI

PCK and CK remain among the most important components of teaching. It also contributed to the skeleton of the TPACK framework, which focuses on how teachers use technology, such as AI, to teach. The key aspect of this review was how the use of GAI influenced chemistry teachers PCK and CK. The teacher would be able to demonstrate this through dialogue with the AI chatbot. Through the interaction, teachers were able to ask the chatbot for assistance in addressing the misconception about the -OH group for a base and the one for the alcohol, where consequently the responses were not what the teacher was anticipating. However, the teachers were able to adapt their prompts until the chatbot provided a satisfactory answer, even though it was insufficient. The teacher could not verify the references provided by the chatbots. This highlights the significance of sufficient PCK and CK when teachers use GAI for lesson planning, as their absence is likely to bring along misconceptions. GAI chatbots seem to be of less significance in addressing students' confusion regarding representation levels when learning chemistry. For instance, it would provide a definition of a certain type of bond (metallic, ionic, covalent bonds) using a particular representation while using another representation to provide a definition of another type of bond. This may create difficulties for teachers who use GAI for lesson planning, as they will have to explain different bonds in different representations, resulting in more confusion and misconceptions during their teaching of chemistry.

Another key aspect of teaching chemistry is the visualization of chemistry content, which has massive consequences for the understanding of chemistry, a subject that is highly characterized by the molecular and particle nature of matter. Furthermore, ChatGPT 4.0 demonstrated tutoring abilities when students engaged with it. The students were able to understand complex figures, analyze chemical structures, interpret tables, and analyze graphs. Nevertheless, the chatbot could not assist when the diagrams were disoriented or of poor quality. Chatbots also have the ability to generate images after being prompted through text. However, this was applicable to ChatGPT 4.0, despite minimal errors, while its counterparts, such as Copilot/BingChat and Google Bard/Gemini, seemed to underperform on that front. Furthermore, Google Bard/Gemini was found to be more wanting, as it struggled to identify a substrate in a reaction. Moreover, GAI struggles when it is expected to represent covalent bonding, where it misplaces and misrepresents the position and orientation of atomic particles. These misrepresentations have dire implications for the conceptual understanding of chemistry during the teaching process if teachers decide to rely on them without cross-checking with other sources. It is clear that, through its hallucinations, the GAI chatbots would use the macro level to represent sub-microscopic aspects such as ionic bonds. However, the use of Python coding instead of DALL-e during the prompting strategy enhanced the responses generated by the chatbot. For example, when the prompting involved the generation of Lewis structures. Furthermore, another positive aspect of GAI is its ability to produce text. This is expected from a large language model that is trained more on producing and refining text. Nonetheless, the chatbots seem to provide incorrect definitions, such as those of covalent bonding. The chatbot seems to demonstrate poor content knowledge, which may result in misconceptions if the teacher uses them without confirming from other sources. The other challenge observed was that the chatbots personified atoms, again with implications for misconceptions. The final aspect that is important for the teaching of chemistry is the approach the teacher uses to teach. PBL is a strategy that aligns well with the nature of science, and its use cannot be easily disregarded. Teachers may use chatbots to design PBL tasks. So does students during the conceptualisation, data collection or any step. They can also use the chatbot to suggest topics and even suggest additional topics or subtopics for their group members. Students can use ChatGPT, for example, to simulate drying curves and particle distributions. Nonetheless, the level of PCK should be at the top, which may require students to continuously consult with their teachers throughout the process.

Limitations

This study focused on the use of GAI in chemistry education and did not involve other science education subjects such as physics education and life sciences education. Moreover, the review focused on the teaching and learning aspects of chemistry, but could not go deeply into other aspects such as motivation, self-directed learning, self-regulation, achievements, or technology acceptance, which also have positive implications for the learning process. Furthermore, artificial intelligence is a very broad area that may include subareas such as machine learning, deep learning, and robotics. This study focused only on GAI. Above all, the papers included in the review did not focus on all types of GAI chatbots. Only ChatGPT 3.5, ChatGPT 4.0, Google Bard, Bing Chat, Adobe Firefly, Leonardo AI, DALLE-E were used in the reviewed studies. Furthermore, improved versions of the indicated chatbots currently exist, which did not exist when the reviewed studies were undertaken.

Conclusion and Future Recommendations

This paper provides an overview of the use of GAI in teaching and learning chemistry. The review noted that, with the use of the GAI, students were able to demonstrate critical thinking by posing questions, evaluating the responses from the chatbots, and consolidating the information through summaries. Through this, students are able to notice the inaccuracy of the information through their assessment of the responses from the chatbot. Furthermore, a chatbot like ChatGPT demonstrates limitations in the problem-solving stage unless the user utilizes a more adapted prompting strategy, such as the 5Cs. Furthermore, students can use AI chatbots for conceptualization and throughout the inquiry process in chemistry education. Nonetheless, students need to be more objective and wary when using ChatGPT for learning. For example, an AI chatbot such as ChatGPT was found wanting in playing the tutoring role, where it showed inconsistencies in the answers that it generated, while on the other hand, it struggled with high order questions –hallucinating in the process. This shortcoming may lead to misconceptions for students who would rely only on the chatbot without verifying information from other sources. Nonetheless, despite struggling with interpreting diagrams that have poor resolution and orientation, ChatGPT 4.0 seems to be a better tutor than its predecessor. Nonetheless, this version is not free and requires a subscription. This may be a challenge for students from disadvantaged backgrounds. Furthermore, the use of GAI has pedagogical implications for chemistry education. The teacher showed elements of TPACK during the prompting stage, such as CK and PCK. While prompting itself relies on the teachers' CK, the teacher would use their PCK during their dialogue with ChatGPT in asking for the difference between ionic and molecular materials while also seeking relevant teaching strategies, including those that eliminate a misconception on the 'OH' from NaOH and CH3CH2OH. After noting the inability of the chatbot to provide satisfactory responses, the teacher modified their prompts. Another aspect of chemistry teaching is the representation and visualization of chemistry using GAI. The free version of ChatGPT performed well in interpreting diagrams but struggled with poor-quality images, whereas ChatGPT 4.0 was able to read tables and graphs, student handwritten texts, and spectra analysis.

Recommendations

However, future research should explore this further. When used effectively, teachers can integrate GAI into their teaching. Another point to note is that GAI chatbots produce the desired responses when users employ Python code compared to DALL-e. Furthermore, GAI chatbots struggle with representation levels, where they provide correct-looking responses with incorrect representation. This may require a more knowledgeable user. Moreover, although GAIs does well in producing texts, it personifies particles during definitions. This may lead students to view particles as living beings, creating more misconceptions. Lastly, the chatbots demonstrated the ability to aid in the conceptualization of PBL tasks by suggesting topics that can be assigned to different groups, producing drafts and rubrics, and structuring inquiry stages and preliminary simulations, among others. Nonetheless, what is crucial is that despite all these GAI abilities, the teacher would still have to demonstrate their objectivity, PCK, and CK throughout the prompting stages. Furthermore, teachers should not only rely on the responses from chatbots without verifying them with other sources. Moreover, teachers should not just use the responses from the chatbots as final drafts, but as initial drafts that would still be modified after verification with other sources and the user's own assessment.

Scientific Ethics Declaration

* The author declares that the scientific ethical and legal responsibility of this article published in JESEH journal belongs to the author.

Conflict of Interest

* The author declares that there is no conflict of interest.

Funding

* This research did not receive any funding.

Acknowledgements or Notes

* The author acknowledges the use of software such as Paperpal for language editing.

References

Akaygun, S., & Kilic, I. (2025). Generative Artificial Intelligence (GenAI) as the artist of chemistry visuals: Chemistry preservice teachers' reflections on visuals created by GenAI. *Journal of Chemical Education, 102 (7),* 2549–2564. https://doi.org/10.1021/acs.jchemed.4c00775

Alasadi, E.A., & Baiz, C.R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education, 100* (8), *2965-2971.* https://doi.rg/10.1021/acs.jchemed.3c00323

Binkley ,M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In Griffin, P., McGraw, B., & Care, E. (Eds). *Assessment and teaching of the 21st century skills*. New York: Springer

Buthelezi, S. (2025, April 14). Ethical debate surrounding AI use by students in South African universities. *Independent Online (IOL).* https://iol.co.za/mercury/2025-04-14-ethical-debate-surrounding-ai-use-by-students-in-south-african-universities/

Clark, T. M. (2023). Investigating the use of an artificial intelligence chatbot with general chemistry exam questions. *Journal of Chemical Education*, *100*(5), 1905–1916. https://doi.org/10.1021/acs.jchemed.3c00027

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

Elmas, R., Adiguzel-Ulutas, M., & Yılmaz, M. (2024). Examining ChatGPT's validity as a source for scientific inquiry and its misconceptions regarding cell energy metabolism. *Education and Information Technologies*, *29*(18), 25427–25456. https://doi.org/10.1007/s10639-024-12749-1

Fenta, A. A. (2025). A review on enhancing education with AI: exploring the potential of ChatGPT, Bard, and generative AI. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00426-5

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education*, *100*(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

Haleem, A., Javaid, M., & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *2*(4). https://doi.org/10.1016/j.tbench.2023.100089

Haryani, E., Cobern, W. W., Pleasants, B. A-S. & Fetters, M. (2024). Exploring pedagogical strategies: Integrating 21st-century skills in science classrooms. *Journal of Education in Science, Enviroment and Health, 10*(2), 106-119. https://doi.org/10.55549/jeseh.697

Iyamuremye, A., Niyonzima, F. N., Mukiza, J., Twagilimana, I., Nyirahabimana, P., Nsengimana, T., Habiyaremye, J. D., Habimana, O., & Nsabayezu, E. (2024). Utilization of artificial intelligence and machine learning in chemistry education: a critical review. In *Discover Education, 3*(1). https://doi.org/10.1007/s44217-024-00197-5

Kim, J. (2025). Integrating Artificial Intelligence (AI) chatbots and green chemistry principles in the synthesis of cyclohexene. *Journal of Chemical Education, 102*, 3058-3064. https://doi.org/10.1021/acs.jchemed.5c00212

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

Lawasi, M. C., Rohman, V. A., & Shoreamanis, M. (2024). The use of AI in improving student's critical thinking skills. *Proceedings Series on Social Sciences & Humanities*, *18*, 366–370. https://doi.org/10.30595/pssh.v18i.1279

Ling Jen, S., & Rahim Hj Salam, A. (2024). Using Google bard to improve secondary school students' essay writing performance. *Journal of Creative Practices in Language Learning and Teaching, 12* (1). https://doi.org/10.24191/cplt.v12i1.24999

Leon, A.J., & Vidhani, D. (2023). ChatGPT needs a chemistry tutor too. *Journal of Chemical Education,100 (10)*, 3859–3865. https://doi.org/10.1021/acs.jchemed.3c00288

Lorenz, U., Romeike, R. (2023). What is AI-PACK? – Outline of AI competencies for teaching with DPACK. In: Pellet, JP., Parriaux, G. (eds) Informatics in schools. beyond bits and bytes: Nurturing informatics intelligence in education. ISSEP 2023. *Lecture Notes in Computer Science*, *14296.* Springer, Cham. https://doi.org/10.1007/978-3-031-44900-0_2

Murni, H.P.,Azhar, M.,Ellizar, E., Nizar, U. K. & Guspatni, G. (2022). Three levels of chemical representation-integrated and structured inquiry-based reaction rate module: its effect on students' mental models. *Journal of Turkish Science Education, 19(3),* 758-772. https://doi.org/10.36681/tused.2022.148

Nascimento Júnior, W. J. D., Morais, C., & Girotto Júnior, G. (2024). Enhancing AI responses in chemistry: Integrating text generation, image creation, and image interpretation through different levels of prompts. *Journal of Chemical Education, 101*(9), 3767–3779. https://doi.org/10.1021/acs.jchemed.4c00230

Pabuçcu-Akış, A. (2024). Using innovative technology tools in organic chemistry education: bibliometric analysis. *Chemistry Teacher International, 7*, 141 - 156. https://doi.org/10.1515/cti-2024-0055

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review, 14,* 47–61. https://doi.org/10.1016/j.edurev.2015.02.003

Ramos, B., & Condotta, R. (2024). Enhancing learning and collaboration in a unit operations course: Using AI as a catalyst to create engaging problem-based learning scenarios. *Journal of Chemical Education, 101(8),* 3246–3254. https://doi.org/10.1021/acs.jchemed.4c00244

Ruff, E. F., Engen, M. A., Franz, J. L., Mauser, J. F., West, J. K., & Zemke, J. M. (2024). ChatGPT writing assistance and evaluation assignments across the chemistry curriculum. *Journal of Chemical Education, 101(6)*, 2483-2492. https://doi.org/10.1021/acs.jchemed.4c00248

Sedagat, S. (2025). Plagiarism and wrong content as potential challenges of using chatbots like ChatGPT in medical research. *Journal of Academic Ethics, 23*, 185-188. https://doi.org/10.1007/s10805-024-09533-8

Sykes, D. (2023). Exploring the use of ChatGPT in lesson planning: Possibilities, experiences, and limitations. *Literacies and Language Education: Research and Practice*, 39-51. English Language Institute, KUIS.

Talanquer, V. (2011). Macro, Submicro, and Symbolic: The many faces of the chemistry "triplet". *International Journal of Science Eduction, 33*:2, 179-195. https://doi.org/10.1080/09500690903386435

Tassoti, S. (2024). Assessment of students use of generative artificial intelligence: Prompting strategies and prompt engineering in chemistry education. *Journal of Chemical Education*, *101*(6), 2475–2482. https://doi.org/10.1021/acs.jchemed.4c00212

Valeri, F., Nilsson, P., & Cederqvist, A. M. (2025). Exploring students' experience of ChatGPT in STEM education. *Computer and Education: Artificial Intelligence,* https://doi.org/10.1016/j.caeai.2024.100360

Wangdi, T., Rigdel, S.K., Dawa, T., & Tshering, T. (2025). Using ChatGPT as an assessment tool in education: A systematic literature review of practices and limitations. *Issues in Educational Research, 35(2),* 818-837. http://www.iier.org.au/iier35/wangdi.pdf

Author(s) Information

Tebogo E. Nkanyani
University of Pretoria,
George Storrar Dr &, Leyds St, Groenkloof, Pretoria, 0027,
South Africa
Contact e-mail: tebogo.nkanyani@*up.ac.za*
ORCID iD: 0000-0003-4924-3882