Ranah Research
Journal of Multidisciplinary Research and Development
E-ISSN: 2655-0865
082170743613 ranahresearch@gmail.com https://jurnal.ranahresearch.com
DINASTI RESEARCH

DOI: https://doi.org/10.38035/rrj.v8i5

Utilization of AI Chat Bot as a Chemistry Learning System on Atomic Structure to Improve Students' Understanding of Concepts: Systematic Review

Bunga Adinda Putri[1], Isti Yunita[2]
[1]Universitas Negeri Yogyakarta, Yogyakarta, Indonesia,
[2]Universitas Negeri Yogyakarta, Yogyakarta, Indonesia,

Corresponding Author: bungaadinda.2025@student.uny.ac.id[1]

Abstract: This study aims to analyze the role and implementation of artificial intelligence (AI)-based chatbots in chemistry learning at the high school level, particularly in strengthening understanding of abstract concepts such as atomic structure. This study uses a systematic literature review approach to 18 relevant scientific articles from 2016 to 2026 originating from reputable international journals, preprints, and empirical research results in the fields of educational technology and chemistry education. The results of the study indicate that the use of AI chatbots in high school chemistry learning can be grouped into two main categories: studies that focus on the development of AI-based learning systems and studies that emphasize the implementation of chatbots as learning aids such as virtual tutors, concept exploration agents, simulation media, and scientific writing support. In general, AI chatbots have been proven to improve student learning motivation, understanding of chemical concepts, critical thinking skills, and independent learning skills. However, several limitations were also found, such as inaccurate information, hallucination phenomena, lack of adaptation to the high school pedagogical context, and the risk of student dependence on AI technology. Specifically for atomic structure, there has been no research specifically and in-depth on the implementation of AI chatbots, leaving a significant research gap. Therefore, the development of a more contextual, targeted, and tailored AI chatbot-based chemistry learning model is needed, tailored to the characteristics of high school students.

Keyword: AI Chatbot, High School Chemistry Learning, Atomic Structure, Artificial Intelligence, Systematic Literature Review, Educational Technology.

INTRODUCTION

Digital transformation in education is experiencing significant acceleration along with the development of technology. *Artificial Intelligence* (AI), particularly in supporting adaptive and personalized learning, has emerged as a transformative force in educational settings. One innovation that has been widely studied is the use of AI-based chatbots, including *large language models* such as ChatGPT, which is capable of providing interactive and contextual responses to users. This technology enables the creation of more flexible, interactive, and student-centered learning (*personalized learning*), thus potentially increasing the effectiveness of the learning process (Kasneci et al., 2023; Zhai, 2023). Furthermore, systematic reviews show that AI chatbots can improve learning accessibility and provide real-time feedback to students (Labadze et al., 2023).

In chemistry learning, one of the main challenges is students' poor conceptual understanding due to the abstract and complex nature of the material. The fundamental concept of atomic structure in chemistry is often difficult to grasp because it cannot be directly observed and requires strong visualization skills. Recent research shows that students still struggle to grasp chemical concepts in depth, especially when learning is not supported by an approach that aligns with the material's characteristics (Taber, 2021). Furthermore, this difficulty is also related to students' limited ability to connect comprehensive representations of chemical concepts, such as macroscopic, submicroscopic, and symbolic levels, which are essential for understanding atomic structure.

This problem is further amplified by the use of conventional learning methods which are still predominantly teacher-centered (*teacher-centered learning*). This approach tends to limit active student engagement and provides less direct feedback on their understanding. Studies show that less interactive learning can hinder students' knowledge construction and result in poor conceptual understanding (Holmes et al., 2022). Therefore, learning innovations are needed that can increase student interaction, engagement, and learning independence, particularly for abstract topics such as atomic structure.

The use of AI-based chatbots is a relevant alternative solution to address these challenges. Chatbots have the ability to provide instant responses, answer student questions, and provide explanations tailored to learning needs. Literature reviews have shown that chatbots can increase student engagement, support self-paced learning, and provide a more interactive learning experience than traditional methods (Wollny et al., 2021; Okonkwo & Ade-Ibijola, 2021). Furthermore, chatbots can also function as virtual learning assistants, helping students understand material more deeply through dialogic interactions. In the context of atomic structure, these capabilities are particularly valuable, as students often require repeated explanations and varied representational formats to grasp abstract concepts such as electron configuration, quantum numbers, and orbital theory.

In the context of science learning, the integration of AI technology, including chatbots, has demonstrated a positive impact on student learning outcomes. This technology enables more adaptive material delivery and supports the development of critical thinking and problem-solving skills. AI-based chatbots can also help explain complex concepts through simpler and more contextual approaches, making it easier for students to understand abstract material (Kasneci et al., 2023). Thus, the use of chatbots in chemistry learning has significant potential to improve the quality of learning, particularly in understanding the concept of atomic structure.

However, studies specifically addressing the use of AI chatbots in chemistry learning, particularly in atomic structure, are still limited. Most studies focus on the use of chatbots in general education, thus not providing a comprehensive overview of their effectiveness in the context of chemistry learning (Labadze et al., 2023). Moreover, preliminary observations of the existing literature indicate that direct empirical investigations into AI chatbot implementation specifically for atomic structure topics at the secondary school level remain scarce. This scarcity necessitates a broader analytical approach to derive meaningful insights.

Given this gap, the present study adopts a systematic literature review (*Systematic Literature Review*) to identify, analyze, and synthesize relevant research results from broader chemistry and STEM education contexts. Recognizing the limited number of primary studies specifically targeting atomic structure, this review strategically synthesizes findings from the broader chemistry and STEM education literature to derive applicable frameworks and pedagogical implications for atomic structure learning. By doing so, this study aims to: (1)

map the existing forms and implementations of AI chatbots in chemistry education, (2) identify the characteristics of AI chatbot-based learning systems relevant to abstract chemical concepts, (3) analyze the advantages and limitations of AI chatbot integration, and (4) evaluate the effectiveness of AI chatbots in improving conceptual understanding while highlighting specific research directions for atomic structure.

Based on this description, this study aims to systematically examine the use of AI chatbots as a chemistry learning system for atomic structure to improve students' conceptual understanding. The results of this study are expected to contribute to the development of technology-based learning innovations and serve as an empirical basis for implementing AI chatbots in chemistry learning at the secondary school level (Zhai, 2023), while also identifying critical research directions for future investigations specifically focused on atomic structure.

METHOD

Research Design

This study uses the Systematic Literature Review (SLR) method to identify, evaluate, and synthesize research related to the use of web-based chatbots.*Artificial Intelligence*(AI) in chemistry learning, particularly on atomic structure, to improve students' conceptual understanding. The SLR method was chosen because it can provide a comprehensive understanding of research developments in a field through a systematic, transparent, and replicable process (Labadze et al., 2023).

Research Questions (RQ)

Based on the research objectives, the research questions in this study are formulated as follows:

a) RQ1: What is the form and implementation of AI chatbot-based learning in chemistry education?
b) RQ2: What are the characteristics of an AI chatbot-based learning system on atomic structure material?
c) RQ3: What are the advantages and limitations of using AI chatbots in chemistry learning?
d) RQ4: How effective is the use of AI chatbots in improving students' conceptual understanding in chemistry learning?

PICO Framework

To clarify the scope of the research and assist the literature search process, this research uses the PICO (Population, Intervention, Comparison, Outcome) framework.

Table 1. PICO Framework

Component	Description
Population (P)	Students in chemistry learning (high school/equivalent level)
Intervention (I)	AI chatbot-based learning systems (e.g. ChatGPT or*conversational AI*)
Comparison (C)	Conventional learning methods or non-AI learning media
Outcome (O)	Improved conceptual understanding, learning outcomes, and student engagement

The use of the PICO framework helps ensure that the literature search process remains focused on research relevant to the topic of using AI chatbots in chemistry learning.

Literature Search Strategy

The literature search process was carried out systematically using three international scientific databases, namely*Scopus*, *ScienceDirect* And *ERIC* These three databases were selected because they offer a broad and high-quality range of publications in education, science, and technology. The use of multiple databases aims to improve the completeness and reliability of the literature search results. The publication year range used was 2016–2025, ensuring that the articles analyzed reflect the latest developments related to the use of AI chatbots in learning.

Search keywords are arranged based on PICO components and tailored to the research topic.*search string*used are:

("AI chatbot" OR "chatbot" OR "artificial intelligence" OR "ChatGPT" OR "conversational AI" OR "large language model" OR "LLM") AND ("chemistry education" OR "chemistry learning" OR "science education" OR "STEM education") AND ("conceptual understanding" OR "learning outcomes" OR "student understanding" OR "academic achievement")

This search string is applied to each database with syntax adjustments according to the search system used.

Inclusion and Exclusion Criteria

In this study, the selection of articles for analysis was carried out through the application of inclusion and exclusion criteria to ensure the relevance and quality of the literature used. The inclusion criteria included several things, namely articles published between 2016 and 2025, thus representing the latest research developments related to the use of chatbots.*Artificial Intelligence*(AI) in learning. Furthermore, selected articles must be indexed in reputable databases, such as Scopus, ScienceDirect, or ERIC, to ensure the quality and credibility of the sources used.

Furthermore, articles included in this study must be scientific publications such as journals, proceedings, or other reputable scientific works. The focus of the article should also relate to the use of AI chatbots in educational contexts, particularly those relevant to chemistry or science learning. Furthermore, the article should examine aspects of conceptual understanding, learning outcomes, or learning effectiveness, thus aligning with the objectives of this study.

Meanwhile, exclusion criteria were applied to filter out articles that did not align with the research focus. Articles were excluded if they were not relevant to the use of chatbots or AI in education and did not discuss chemistry or science learning. Furthermore, non-scholarly publications such as blogs, opinion pieces, and editorials were also excluded from this research analysis because they did not meet academic standards. Articles not available in full text (*full text*) were also excluded because they did not allow for in-depth analysis. Finally, articles identified as duplicates from other databases were eliminated to avoid data redundancy in the analysis process.

Study Selection Process

The literature selection process in this study was conducted in accordance with the PRISMA guidelines to ensure transparency, consistency, and replicability in the article selection process. This approach allowed researchers to systematically screen the literature so that only relevant and high-quality articles were used in the analysis.

The selection stage begins with the stage*identification*, namely the process of identifying all articles obtained from the database used through the application of keywords (*search string*) that have been previously determined. At this stage, all relevant articles are collected without considering duplication or in-depth suitability.

Next, at the stage*screening*An initial screening process was carried out by removing duplicate articles and selecting articles based on their titles and abstracts. This stage aimed to eliminate articles that clearly did not align with the research focus.

The next stage is*eligibility*, namely the assessment of the suitability of an article by reading the full text (*full text*). At this stage, articles are evaluated in more depth based on the established inclusion and exclusion criteria to ensure they are appropriate to the research objectives.

The last stage is*included*, namely determining the final articles that meet all criteria and are worthy of further analysis in this study. Articles that pass this stage are then used as the basis for the data synthesis and analysis process.

Data Extraction

Data from each selected article was collected through a systematic data extraction process to obtain information relevant to the research objectives. The extracted data included the author's name, year of publication, research objective, type of AI chatbot used, learning context (specifically atomic structure), research method, subjects or samples, and key research findings, including effectiveness, strengths, and limitations. This data was then used as the basis for analysis and comparison between studies.

Data analysis

The data obtained were analyzed using qualitative synthesis and comparative analysis methods. The analysis process was carried out by grouping studies by theme, identifying patterns of AI chatbot use in learning, comparing effectiveness between studies, and analyzing the strengths and limitations of each study. The results of this analysis were used to identify research trends, research gaps, and the effectiveness of AI chatbot use in improving students' conceptual understanding in chemistry learning.

RESULTS AND DISCUSSION

The literature selection process in this study was conducted in accordance with the PRISMA guidelines to ensure transparency, consistency, and replicability, consisting of four main stages: identification, screening, eligibility, and inclusion. In the identification stage, all articles were retrieved from the Scopus, ScienceDirect, and ERIC databases using the predefined search string, yielding a total of 1,857 documents. Subsequently, in the screening stage, 1,483 duplicate records were removed, leaving 374 articles, which were then screened based on their titles and abstracts, resulting in the exclusion of 356 articles because they did not align with the subject focus (e.g., non-educational technology or non-chemistry), were not in English, were published before 2016, or were not journal articles, thus leaving 18 articles to proceed to the eligibility stage. In the eligibility assessment, the full texts of the remaining 18 articles were thoroughly reviewed and all were confirmed to meet the inclusion criteria and deemed relevant for a comprehensive analysis. Finally, in the inclusion stage, all 18 articles were selected for the final analysis and data synthesis in this study. The final results of this process are then presented in the form of a study selection flowchart as shown in Figure 1 (PRISMA Diagram).

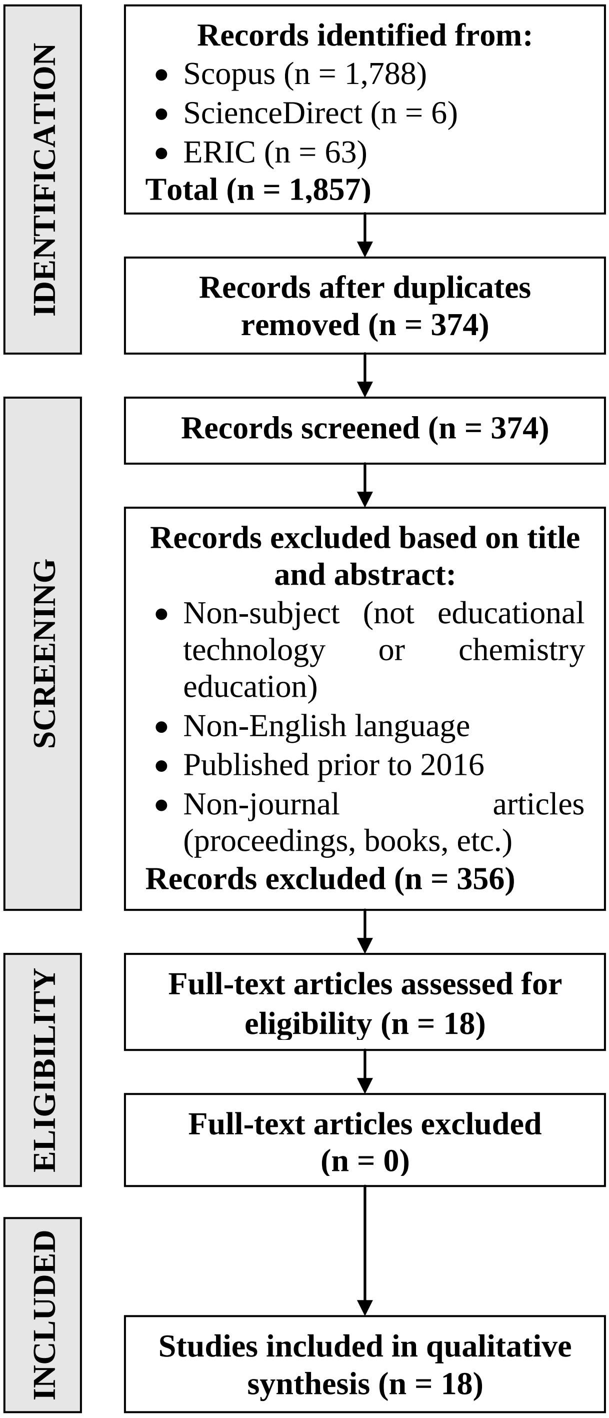

Figure 1. PRISMA Diagram

Characteristics of Articles

Table 2. Article Characteristics (n:18)

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
1	Morbidoni et al. (2026)	Evaluative study (quantitative-based comparative performance analysis)	High school and university level chemistry questions (MCQ & essay: 150 high school questions, 75 university questions, 100 stoichiometry questions)	Comparison of ChatGPT model performance (GPT-3.5 Turbo, GPT-4 Turbo, GPT-4o, o1-preview, GPT-5) in solving	The latest GPT models (o1-preview and GPT-5) show the highest accuracy and reasoning ability; computational and interpretation errors are still found, especially in redox reactions and chemical equilibria.

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
				chemistry problems	
2	Cheung (2026)	Systematic Literature Review (SLR) on 19 empirical studies	19 empirical articles on AI chatbots in science education	AI chatbot in science learning based on science identity (interest, competence, recognition, epistemic affect)	AI chatbots have mostly focused on increasing students' interest in science, but are still lacking in developing epistemic affect, scientific competence, and recognition of identity as a "science person"; it is recommended to use a more comprehensive science identity model in learning design.
3	Gill et al. (2023)	Conceptual review + expert discussion (qualitative synthesis)	Literature review and conceptual analysis across educational fields (finance, coding, mathematics, etc.)	ChatGPT as an AI chatbot in modern education	ChatGPT has great potential in supporting learning (material creation, tutoring, and discussions), but has limitations such as hallucinations, information inaccuracies, the risk of plagiarism, and is not yet able to guarantee high-validity communication; AI regulations and literacy are needed for educators and students.
4	Lee & Zhai (2023)	Content analysis + lesson plan evaluation based on TPACK rubric	29 pre-service elementary teachers (universities in Korea)	Using ChatGPT in preparing science lesson plans	Pre-service teachers were able to integrate ChatGPT into science lessons (Biology, Chemistry, and Earth Science). The average TPACK score was high (M = 3.29/4), demonstrating ChatGPT's potential to support independent learning, questioning, and formative assessment. However, limitations were identified in optimizing ChatGPT's functionality, as well as concerns regarding information accuracy and student dependency.
5	McAfee & Rodriguez (2025)	Qualitative research comparison (AI-assisted coding vs human analysis), studi eksploratif	Academic integrity analysis dataset on chemistry syllabus (secondary data from previous research)	ChatGPT as a “research assistant” in qualitative data analysis of chemistry education	ChatGPT is capable of coding qualitative data and providing analyses comparable to those of human researchers, including inter-rater reliability evaluations. However, the results suggest that AI should remain positioned as a "research partner," not a replacement for researchers, as transparency, validation,

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
					and methodological control are essential in research.
6	Babalola et al. (2024)	Correlational	50 high school students	ChatGPT in organic chemistry learning	There is a positive relationship between the social environment and students' interests, critical thinking, and learning achievements in using ChatGPT.
7	Ruff, Franz, & West (2024)	Implementation of learning + evaluation of laboratory assignments + student surveys	Advanced chemistry students (analytical & inorganic chemistry upper-level laboratory courses)	ChatGPT as an aid to method development and green chemistry in laboratory proposals	ChatGPT provided initial ideas but tended to produce general and inaccurate answers, especially for advanced methods and green chemistry. Students were still able to critically evaluate the AI's output, but confidence in the AI was low due to a lack of references. This activity remains effective for practicing critical thinking and evaluating AI in the context of chemistry.
8	dos Santos (2024)	Single-case study (learning interaction log analysis)	Students in a chemistry learning simulation	ChatGPT and Bing Chat as*agents-to-think-with*in learning chemistry	ChatGPT and Bing Chat can improve critical thinking, problem-solving, conceptual understanding, creativity, and personalized learning. ChatGPT performed better than Bing Chat in terms of depth, completeness, and accuracy of answers. Effective use requires prompt crafting skills and teacher training for AI integration in learning.
9	Rojas (2024)	Implementation of learning + analysis of student experience	College level chemistry students (inorganic chemistry class)	ChatGPT as a tool for scientific writing assignments	ChatGPT was effective in ideation, topic verification, proofreading, and editing. Students reported positive experiences and found ChatGPT helpful in improving the quality of their writing, although ethical issues remain regarding the use of AI in academic writing.
10	Chan et al. (2024)	Comparative experiment + expert evaluation (subject-matter experts)	Pre-university level STEM questions (chemistry, physics, mathematics)	ChatGPT (GPT-3.5 & GPT-4) for automatic item generation via prompt engineering	LLM was able to generate high-quality questions, especially with chain-of-thought prompting, but still had errors in multi-step steps and the risk of hallucinations. GPT-4 with coding + chain-of-thought

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
					produced the best performance in terms of final answer accuracy and STEM question quality.
11	Rücker & Becker-Genschow (2025)	Cluster randomized controlled trial (pre-post test)	195 9th grade students of secondary schools in Germany	Domain-specific AI chatbot (ADA) as a personalized learning assistant in STEM mathematics learning	Chatbots increased student motivation, situational interest, and learning outcomes. There was high acceptance of the technology. However, cognitive load increased slightly. Chatbots were effective as personalized learning support and instructional differentiation.
12	Lolinco & Holme (2023)	Development study + large class implementation + chatbot interaction analysis	347 early-level chemistry students (large lecture class)	Curation chatbot as an exploratory communication tool and aid in writing chemical contexts	Chatbots are widely used (78% of students) to assist with idea development and chemistry essay writing. The majority of interactions are used to aid in assignment ideation. Chatbots are effective for contextual learning, but serve more as a tool to support idea exploration than as a primary source of material.
13	Davidovitch & Gerkerova (2026)	Correlational quantitative survey	76 students from 5 universities in Ukraine (various fields: STEM, humanities, social)	AI in higher education (including chatbots, adaptive learning, virtual assistants)	AI use is positively associated with motivation, engagement, and academic achievement. STEM students have the most positive attitudes. However, ethical challenges such as plagiarism, data privacy, and the AI literacy gap exist, requiring regulation and training.
14	Shishakly (2025)	Quantitative survey (structured questionnaire)	506 students from several higher education institutions in the UAE	ChatGPT in academic learning and writing	There are differences in ChatGPT usage based on gender, level of study, and field of study. Female students and final-year students use ChatGPT more frequently. ChatGPT improves academic writing skills, especially for non-native English speakers, but training is needed for more equitable and effective use.
15	Lee & Song (2024)	Mixed-method (perception experiment + quantitative	11 teachers and 70 elementary school students (grade 6)	ChatGPT as a concept explanation generator in	Teachers and students have different perceptions of AI explanations. Teachers find AI effective

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
		and qualitative analysis)		computer learning (GAI-generated concept explanations)	for certain concepts, while students struggle to distinguish between explanation sources. AI is perceived as enhancing clarity and readability, but better AI literacy and pedagogical design are needed for classroom integration.
16	Chen, Chen, & Huang (2026)	Quasi-experiment (experimental vs control group) + usability questionnaire	Chemistry students/participants in a virtual 3D laboratory	Enhanced AI Chatbot (EAIC) based on ChatGPT + Rasa as a virtual tutor in a 3D chemistry laboratory	EAIC increased learning satisfaction, technology acceptance, and perceived usefulness. There was no significant difference in learning performance, but students were more satisfied with the chatbot interaction. Weaknesses were found in accuracy, conversational context, and dialogue clarity, necessitating improvements to the conversational AI system.
17	Rababa & Al-Omari (2025)	Quantitative (test of mastery of scientific concepts)	396 elementary school students (grade 10) in Irbid	Teachers' use of AI in science learning (not specifically chatbots, more towards educational AI applications)	Moderate use of AI results in better mastery of scientific concepts than low and high levels; excessive use of AI is not always pedagogically effective.
18	Lee et al. (2024)	Randomized Controlled Trial (RCT) dalam blended learning	61 undergraduate students in basic chemistry courses in Taiwan	The use of ChatGPT with guidance mechanisms (Guidance-based ChatGPT / GCLA) to support self-regulated learning and higher-order thinking skills	Guided ChatGPT (giving hints, not direct answers) improves SRL, HOTS, and knowledge construction compared to regular ChatGPT; guided use is more effective than free use.

Based on the 18 articles analyzed in this study, the literature was then grouped into two main categories: core articles and supporting articles. This was based on the limited research specifically addressing the application of AI chatbots in high school chemistry education, particularly in atomic structure. Of the 18 articles analyzed, most studies focused on broader contexts such as higher education, STEM in general, or the use of AI in learning without specific chemistry topics.

Main Studies

Table 3. Main Studies

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
1	Lee et al. (2024)	RCT blended learning	61 chemistry students (Taiwan)	Guidance-based ChatGPT (GCLA)	Improved SRL, HOTS, and knowledge construction compared to regular ChatGPT
2	Chen et al. (2026)	Quasi-experiment + usability test	3D chemistry lab students	EAIC (ChatGPT + Virtual tutor experience)	Increase satisfaction & acceptance, not significant on learning performance
3	dos Santos (2024)	Single-case study	Chemistry simulation students	ChatGPT & Bing Chat agents-to-think-with	Improve critical thinking & problem solving; ChatGPT is superior
4	Chan et al. (2024)	Experiment + expert evaluation	Pre-university STEM questions	LLM prompting GPT-3.5/4	Chain-of-thought improves problem quality; GPT-4 is most accurate
5	Rücker & Becker-Genschow (2025)	Cluster RCT	195 junior high school students	Domain-specific chatbot (ADA)	Increase motivation, interest, and learning outcomes
6	Lolinco & Holme (2023)	Development + classroom study	347 chemistry students	Curated chatbot writing assistant	Helps ideation & writing, widely used by students
7	Rojas (2024)	Classroom implementation study	Chemistry student	ChatGPT writing assistant	Assisting with ideation, editing, proofreading; ethical issues arise
8	Lee & Zhai (2023)	Content analysis + TPACK rubric	29 pre-service teachers	ChatGPT lesson planning	Good TPACK integration; helps inquiry & assessment
9	Lee et al. (2024)	Mixed-method experimental perception	11 teachers & 70 students	GAI concept explanation (ChatGPT)	AI helps clarity, but students have difficulty distinguishing sources
10	Babalola et al. (2024)	Correlational	50 high school students	ChatGPT organic chemistry learning	Social environment influences interest & achievement
11	Lee et al. (2024)	RCT blended learning	61 students	Guidance ChatGPT (GCLA)	SRL & HOTS improved with guided AI

Supporting Articles

Table 4. Supporting Articles

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
1	Cheung (2026)	Systematic Literature Review	19 studi AI chatbot	Science identity & AI chatbot	The focus is still on interest, less on epistemic affect & competence
2	Gill et al. (2023)	Conceptual review	Educational literature	ChatGPT in education	Great potential but there are hallucinations & academic risks
3	McAfee & Rodriguez (2025)	Qualitative AI-assisted research	Chemistry syllabus dataset	ChatGPT as a research assistant	AI as a partner, not a replacement for researchers
4	Ruff et al. (2024)	Laboratory implementation	Advanced chemistry	ChatGPT green chemistry &	Useful but often inaccurate & too

No	Author & Year	Research methods	Subject / Sample	Focus AI Chatbot	Findings
		study	students	method dev	general
5	Davidovitch & Gerkerova (2026)	Correlational survey	76 multi-university students	AI in higher education	AI improves motivation & performance, but there are ethical issues
6	Shishakly (2025)	Quantitative survey	506 UAE students	ChatGPT adoption behavior	Gender & field of study differences in AI use
7	Rababa & Al-Omari (2025)	Quantitative	396 school students	AI use in science teaching	The medium level of AI usage is the most effective

Discussion

Forms and implementation of AI chatbot-based learning in chemistry education

Based on the 18 articles analyzed, the implementation of AI chatbots in chemistry education shows that this technology has been used in several main forms, namely as virtual tutors, scientific writing aids, simulation-based learning systems, and concept exploration support agents. In their role as virtual tutors, chatbots are used to provide conceptual explanations, feedback, and direct learning guidance to students, as seen in the use of chatbots in 3D virtual chemistry laboratories (Chen et al., 2026) and guide-based learning models such as GCLA, which emphasizes providing step-by-step instructions rather than direct answers (Lee et al., 2024). In this context, when applied to high school atomic structure material, chatbots can function as tutors that explain abstract concepts such as electron configuration and atomic orbitals in a step-by-step manner according to the student's level of understanding.

Furthermore, chatbots are also widely used as scientific writing aids, with students utilizing them for ideation, drafting, and revising academic papers. This is demonstrated in studies by Rojas (2024) and Ruff et al. (2024), who found that ChatGPT aids the writing process, although it still has limitations in content depth. In the context of atomic structure, this function is relevant for helping students compile reports on the development of atomic models or explain Bohr's theory and quantum mechanics in scientific writing. Furthermore, chatbots are also being developed as simulation-based learning systems, such as AI integration in virtual laboratories or adaptive learning systems (Chen et al., 2026; Rücker & Becker-Genschow, 2025), which allow students to interact with concepts visually and interactively. In the case of atomic structure, this can be realized in the form of electron distribution simulations or orbital visualizations.

On the other hand, chatbots also function as concept exploration agents or*agents-to-think-with*, a tool that encourages students to think critically through interactive dialogue, as demonstrated by dos Santos (2024) and Lolinco & Holme (2023). In learning atomic structure, this approach allows students to explore conceptual questions such as why electrons don't fall into the nucleus or how the Aufbau principle works through exploratory conversations with AI.

RQ2: Characteristics of AI chatbot-based learning systems on atomic structure material

The main characteristics of AI chatbot-based learning systems in chemistry education indicate that this technology is adaptive, interactive, and personalized. Many studies emphasize that chatbots are able to adjust responses based on students' needs and ability levels, as found in research by Rücker & Becker-Genschow (2025) and Lee et al. (2024), which shows that AI can enhance self-regulated learning and provide a more personalized learning experience. In the context of atomic structure, these characteristics allow students who have not yet grasped basic concepts such as electron configuration to receive simple explanations, while more advanced students can be guided to more complex concepts such as hybrid orbitals.

Furthermore, chatbot systems are interactive because they enable real-time, two-way communication. Studies by dos Santos (2024) and Chen et al. (2026) show that dialogue-based interactions improve conceptual understanding and student engagement. In atomic structure, this allows students to directly ask questions about concepts such as quantum numbers or orbital energy and receive explanations that directly address their needs. Another characteristic is the use of prompt engineering, which affects the quality of AI responses, as demonstrated by Chan et al. (2024) and Morbidoni et al. (2026), where output quality improved when using this approach.*chain-of-thought prompting*This is relevant in learning atomic structure because students can be guided to think step by step in understanding electron configuration. Furthermore, some systems also integrate text-based multimodal learning and visual simulations, such as those found in Chen et al. (2026) and Gill et al. (2023), so that abstract concepts such as atomic orbitals can be visualized more concretely.

Advantages and limitations of using AI chatbots in chemistry learning

The analysis results show that the use of AI chatbots in chemistry learning has several key advantages, particularly in improving students' conceptual understanding, learning motivation, and higher-order thinking skills. Studies by Lee et al. (2024) and dos Santos (2024) show that chatbots can improve higher-order thinking skills, problem-solving abilities, and understanding of scientific concepts. In the context of atomic structure, this is particularly important because the concept is abstract and difficult to understand through verbal explanations alone. Furthermore, chatbots also increase students' learning independence, as found by Shishakly (2025) and Lee et al. (2024), allowing students to learn electron configurations and atomic models independently without always relying on teachers. Chatbots also increase student motivation and engagement (Rücker & Becker-Genschow, 2025; Davidovitch & Gerkerova, 2026) and assist in scientific writing and academic communication (Rojas, 2024; Ruff et al., 2024).

However, there are some significant limitations. One of the main problems is the risk*hallucination*or inaccurate information as found by Gill et al. (2023), Chan et al. (2024), and Morbidoni et al. (2026). In the context of atomic structure, this has the potential to lead to misconceptions such as errors in explaining orbitals or electron configurations. Furthermore, there is the issue of student dependence on AI (Lee et al., 2024; Chen et al., 2026), as well as the limited ability of AI to handle complex, multi-step reasoning (Morbidoni et al., 2026). Lack of AI literacy is also a significant barrier as reported by Shishakly (2025) and Davidovitch & Gerkerova (2026), which shows that not all students are able to use AI optimally and critically.

The effectiveness of using AI chatbots in improving students' conceptual understanding in chemistry learning

Overall, the effectiveness of AI chatbots in chemistry learning can be seen from cognitive, affective, and pedagogical perspectives. From a cognitive perspective, various studies have shown that the use of chatbots can improve students' conceptual understanding, critical thinking skills, and knowledge construction (Lee et al., 2024; Lee & Song, 2024; dos Santos, 2024). In the context of atomic structure, this means students are better able to understand abstract concepts such as electron configuration, orbitals, and quantum numbers through interaction with AI.

From an affective perspective, chatbots have been shown to increase student interest, motivation, and engagement in learning (Rücker & Becker-Genschow, 2025; Davidovitch & Gerkerova, 2026; Babalola et al., 2024; Wollny et al., 2021; Okonkwo & Ade-Ibijola, 2021). This is important because high school students often find atomic structure difficult and abstract. Meanwhile, from a pedagogical perspective, chatbots have proven effective as learning aids when used with a targeted approach, such as using a guide-based system (Lee et al., 2024) or as a tool for exploring ideas (Lolinco & Holme, 2023). However, several studies have also shown that increased understanding does not always directly correlate with improved academic outcomes (Chen et al., 2026), so their effectiveness is highly dependent on the learning design used.

CONCLUSION

Based on the synthesis of 18 analyzed articles, it can be concluded that the use of artificial intelligence-based chatbots in chemistry education has shown significant progress in various implementations, such as virtual tutors, scientific writing aids, simulation-based learning systems, and concept exploration agents. Overall, AI chatbots have proven to have significant potential in supporting the chemistry learning process, particularly in improving conceptual understanding, student engagement, and higher-order thinking skills. However, their effectiveness is greatly influenced by the learning design, the user's AI literacy level, and how the technology is integrated into the learning process (Kasneci et al., 2023; Zhai, 2023; Labadze et al., 2023).

Furthermore, AI chatbots have demonstrated positive contributions in assisting the learning of abstract chemical concepts, such as atomic structure, although most studies have focused on higher education contexts or general science learning. This suggests that the application of chatbots in high school chemistry learning, particularly in atomic structure, has not received much specific research. Therefore, it can be concluded that although AI chatbots have strong potential in supporting chemistry learning, particularly in abstract concepts, there is still a research gap that needs to be filled through more focused studies on atomic structure at the high school level so that their implementation can be evaluated more in-depth and contextually.

Limitations

The main limitation of this study lies in the limited literature specifically discussing the use of AI chatbots in high school chemistry learning, particularly in atomic structure. Of the 18 articles analyzed, most studies focused on higher education contexts, general science learning, or other fields such as mathematics and computer science. Consequently, few directly examined the implementation of AI in atomic structure concepts at the high school level. Consequently, discussions on the application of AI chatbots to this topic remain inferential and developed through generalizations from broader educational contexts.

Furthermore, no research has been found specifically testing the effectiveness of AI chatbots on in-depth understanding of atomic structure concepts, such as students' ability to understand electron configurations, quantum numbers, and quantum mechanical models. Most existing studies only address general chemistry concepts or other topics such as chemical reactions, stoichiometry, and virtual laboratory learning. This situation indicates a significant literature gap between the development of AI research in education and the specific needs of high school chemistry.

Another limitation is the predominance of research that focuses on perception, motivation, or technology use, rather than on in-depth conceptual measurements. Many studies assess chatbot effectiveness in terms of technology acceptance, student engagement, or learning satisfaction, while their impact on mastery of complex chemistry concepts such as atomic structure remains largely unexplored empirically. Consequently, evidence regarding the cognitive effectiveness of chatbots in learning abstract concepts remains limited (Holmes et al., 2022; Taber, 2021).

Recommendation

Based on these limitations, further research is recommended to directly examine the use of AI chatbots in high school chemistry learning on atomic structure, as there are no studies specifically addressing this topic. Research should focus on measuring improvements in students' conceptual understanding, not just motivational or perceptual aspects, to obtain stronger empirical evidence. Furthermore, the development of AI chatbot-based learning models integrated with the curriculum and the improvement of AI literacy for teachers and students are needed to make the use of technology more effective and targeted

REFERENSI

Babalola, V. T., Ahmad, S. S., & Tafida, H. S. (2024). GPT in organic chemistry classrooms: Analyzing the impacts of social environment on students' interest, critical thinking and academic achievement. *Journal of Educational Research in Science Education*.

Chan, C. K. Y., Lee, K. K. W., & Tse, M. C. (2024). Large language models for automatic item generation in STEM education: A comparative study of ChatGPT with prompt engineering. *Computers and Education: Artificial Intelligence, 7*, 100279. https://doi.org/10.1016/j.caeai.2024.100279

Chen, C.-M., Chen, B.-J., & Huang, C.-L. (2026). An enhanced AI chatbot to facilitate learning performance in a virtual 3D chemistry laboratory. *Educational Technology & Society*.

Cheung, K. K. C. (2026). Science identity and AI chatbots: A systematic review of empirical studies in science education. *Journal of Science Education and Technology*. https://doi.org/10.1007/s10956-026-10295-8

Davidovitch, N., & Gerkerova, A. (2026). Artificial intelligence in higher education: Benefits and concerns. *African Educational Research Journal, 14*(1), 10–20.

dos Santos, R. P. (2024). Enhancing chemistry learning with ChatGPT and Bing Chat as agents-to-think-with: A comparative case study. *CIAGE Centre for Generative Artificial Intelligence in Cognition and Education*.

Gill, S. S., Parlikad, A. K., Xu, M., Patros, P., Wu, H., Kaur, R., Kaur, K., Fuller, S., Singh, M., Arora, P., Stankovski, V., Abraham, A., Dustdar, S., Ghosh, S. K., Lutfiyya, H., Kanhere, S. S., Bahsoon, R., Rana, O., Sakellariou, R., ... Buyya, R. (2023). Transformative effects of ChatGPT on modern education: Emerging era of AI chatbots. *arXiv preprint*. https://arxiv.org/abs/2305.12345

Holmes, N. G., Wieman, C. E., & Bonn, D. A. (2022). Teaching critical thinking. *Proceedings of the National Academy of Sciences, 119*(21), e2204017119. https://doi.org/10.1073/pnas.2204017119

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

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), 56. https://doi.org/10.1186/s41239-023-00426-1

Lee, G.-G., & Zhai, X. (2023). Using ChatGPT for science learning: A study on pre-service teachers' lesson planning. *Preprint, University of Georgia*.

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

Lee, S., & Song, K.-S. (2024). Teachers' and students' perceptions of AI-generated concept explanations: Implications for integrating generative AI in computer science education. *Computers and Education: Artificial Intelligence, 7*, 100283.

Lolinco, A. T., & Holme, T. A. (2023). Developing a curated chatbot as an exploratory communication tool for chemistry learning. *Journal of Chemical Education, 100*(10), 4092–4098. https://doi.org/10.1021/acs.jchemed.3c00613

McAfee, S. C., & Rodriguez, J.-M. G. (2025). I asked ChatGPT to do my research: Welcoming artificial intelligence to the chemistry education research team. *Journal of Chemical Education, 102*(12), 5082–5094. https://doi.org/10.1021/acs.jchemed.4c01234

Morbidoni, C., Mascitti, A., Sarra, A., Tonucci, L., & Coccia, F. (2026). Exploring the potential of using ChatGPT in chemistry education. *Journal of Chemical Education, 103*(2), 286–297. https://doi.org/10.1021/acs.jchemed.5c00876

Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence, 2*, 100033. https://doi.org/10.1016/j.caeai.2021.100033

Rababa, H., & Al-Omari, A. (2025). Basic-stage students' acquisition of scientific concepts in light of teachers' use of artificial intelligence applications. *Educational Process: International Journal, 17*, e2025405. https://doi.org/10.22521/edupij.2025.17.405

Rojas, A. J. (2024). An investigation into ChatGPT's application for a scientific writing assignment. *Journal of Chemical Education, 101*(5), 1959–1965. https://doi.org/10.1021/acs.jchemed.3c01289

Rücker, C. R., & Becker-Genschow, S. (2025). Enhancing enthusiasm for STEM education with AI: Domain-specific chatbot as personalized learning assistant. *Computers and Education Open, 9*, 100315. https://doi.org/10.1016/j.caeo.2025.100315

Ruff, E. F., Engen, M. A., Franz, J. L., Mauser, J. F., West, J. K., & Zemke, J. M. O. (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.4c00245

Shishakly, R. (2025). Understanding AI in higher education: Gendered and intersectional students' experiences with ChatGPT use. *European Journal of STEM Education, 10*(1), 36. https://doi.org/10.20897/ejsteme/17646

Taber, K. S. (2021). The nature of student conceptions in science. In *Science Education: An International Course Companion* (pp. 119-131). Brill Sense. https://doi.org/10.1163/9789004505456_009

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet? - A systematic literature review on chatbots in education. *Frontiers in Artificial Intelligence, 4*, 654924. https://doi.org/10.3389/frai.2021.654924

Zhai, X. (2023). ChatGPT and AI in education: A systematic review and future directions. *Journal of Educational Computing Research, 61*(7), 1457-1493. https://doi.org/10.1177/07356331231174823