

## COMPARING TEACHERS AND CHATGPT IN POSING QUALITY QUESTIONS ON MATTER CYCLES

Tufan İnaltekin<sup>1</sup>, Tolga Saka<sup>2</sup>

<sup>1</sup>PhD, Associate Professor, Department of Mathematics & Science, Education, Dede Korkut Faculty of Education, Kafkas University, Türkiye, <https://orcid.org/0000-0002-3843-7393>

<sup>2</sup>PhD, Assistant Professor, Department of Mathematics & Science, Education, Dede Korkut Faculty of Education, Kafkas University, Türkiye. <https://orcid.org/0000-0002-0042-0836>

Corresponding author: [inaltekintufan@gmail.com](mailto:inaltekintufan@gmail.com)<sup>1</sup>  
[tsaka61@gmail.com](mailto:tsaka61@gmail.com)<sup>2</sup>

### Abstract

The purpose of this study is to compare the skills of science teachers and ChatGPT in posing quality questions regarding matter cycles (water, oxygen, carbon, and nitrogen). The study was conducted with data obtained from 34 science teachers and ChatGPT. The research employed a comparative case study design, which is one of the qualitative research approaches. The data from teachers were collected through the Quality Question Posing Form on Matter Cycles, while the data from ChatGPT were obtained using the prompt: "What important questions could you pose to your students about matter cycles?" for each cycle. The data were descriptively analyzed based on the criteria representing the characteristics of quality questions in science education. The findings indicate that ChatGPT met all the quality question criteria across all matter cycles and demonstrated a consistent superiority in aspects such as clarity, alignment with learning objectives, open-endedness, depth, and the absence of guiding elements. In addition, it was observed that most teachers demonstrated competence in the criterion of "avoiding closed-ended questions", while only a few did so in the criteria of "posing probing questions" and "eliciting students' knowledge and understanding". In the nitrogen cycle, fewer teachers were able to pose quality questions compared to the other cycles. In terms of professional experience, teachers with 7–18 years of seniority demonstrated the highest performance in posing quality questions, whereas the  $\leq 3$  years and 19–30 years seniority groups exhibited lower performance. According to the educational level, teachers with a master's degree demonstrated higher performance across all criteria, whereas those with a bachelor's degree were inadequate particularly in posing "deepening questions" and "questions aimed at understanding students". In addition, ChatGPT demonstrated a clear superiority in posing quality questions compared to all educational levels and professional experience groups. The results indicate that ChatGPT was more consistent than teachers in generating effective questions on matter cycles, while teachers' performance varied depending on their professional experience and educational level.

**Keywords:** Science Teacher, ChatGPT, Matter Cycles, Posing Quality Questions

### Introduction

An important part of effective teaching in science is the questions we ask to determine whether students have understood the information presented to them. If, as a teacher, your questions are effective, we can make a sound assessment of the students' understanding and depth of thinking (Ghafar & Hazaymeh, 2024). It can be said that in science, quality questions should include open-ended, curiosity-provoking questions that promote higher-order thinking and align with learning objectives (Bielik & Yarden, 2016). Questioning in the classroom constitutes a central component of dialogue in education (Soysal & Soysal, 2023). Teachers who elicit students' thinking and use it as the basis for instructional decisions can positively influence students' learning (Ruiz-Primo & Furtak, 2007). The structure and cognitive level of the questions that teachers employ in science classes directly affect the quality of students' learning (Talbot et al., 2018). According to Nielsen (2009), a teacher's success in effective teaching is directly related to the use and development of questioning skills. Therefore, posing high-quality questions in science education involves multifaceted and complex pedagogical arrangements and conceptualizations (van Es & Sherin, 2021). Teachers' questioning serves various purposes, including capturing students' attention, guiding thinking, supporting

knowledge construction, and assessing learning objectives (Forster & Penny, 2020; Kayima & Jakobsen, 2020). In science education, oral questions are regarded as fundamental communication tools that promote higher-order scientific thinking among students (Oliveira, 2010). Therefore, posing quality questions helps to bring the authentic spirit of science into education (Vale, 2013). The inclusion of quality questions in science lessons encourages students to engage in deeper and more reflective thinking (Eliasson et al., 2017). Research has shown that teachers' questions play a significant role in promoting dialectical interaction and in fostering students' responsibility for their own learning (Banilower et al., 2013; Chen et al., 2017; Zhai & Tan, 2015). However, according to Wellington and Osborne (2001), many questions posed by science teachers in the classroom are "pseudo-questions" that guide students toward a specific answer. Yip's (2004) study demonstrated that quality thinking questions support students' conceptual change and contribute to their ability to transfer new concepts to different situations. According to Morris and Chi (2020), teachers should reduce questions that direct students toward superficial processes in the classroom and increase quality questions that promote deep and inferential thinking. The types of questions teachers pose and the manner in which they pose them significantly influence the cognitive processes in which students engage while constructing scientific knowledge (Chin, 2007). In science education, "matter cycles" (such as water, carbon, oxygen, and nitrogen) constitute an important learning domain for teachers to help students understand the flow of energy and matter in nature, the interactions between living and non-living entities in a systems-based manner, and to develop scientific thinking skills in this context (Düsing et al., 2019). Matter cycles are a challenging topic for many students, as they involve abstract, complex, and multi-step processes (Barrutia et al., 2021; Sadler et al., 2017; Ugulu et al., 2015). Therefore, posing effective questions in these topics deepens students' conceptual understanding and promotes scientific thinking. In the teaching of such complex and relational topics, the ability to pose quality questions constitutes an indispensable instructional strategy that supports students' higher-order thinking, analytical reasoning, and conceptual meaning-making processes (Idika & Nsifir, 2025; Yip, 2004). The use of artificial intelligence (AI) technologies is increasingly leading to greater scrutiny of teachers' professional knowledge and skills (Almasri, 2024; Park et al., 2023). Moreover, science teachers' questioning skills, like many components of their professional knowledge structures, have been questioned in light of the level reached by AI language models such as ChatGPT. Because AI language models such as ChatGPT have achieved a very high level of content richness in educational settings (Adel et al., 2024). AI-based tools, which are rapidly becoming widespread in education, offer new opportunities to support instructional processes (Rejeb et al., 2024). Advanced AI models, particularly those like ChatGPT, attract attention due to their capacities to provide scientific explanations and generate questions. ChatGPT is capable of generating a wide range of questions on science topics, from simple ones that involve recalling and describing concepts to more complex, higher-order questions that require analysis, synthesis, and inferential thinking (Scaria et al., 2024). Therefore, AI-based tools such as ChatGPT are considered potentially valuable in supporting teachers' instructional processes, for instance in question-generation activities. However, the potential of both sources, teachers and ChatGPT, to generate quality questions on "matter cycles," a challenging topic for middle school students in science, has not yet been investigated. This gap highlights the need to comparatively examine teachers' pedagogical practices alongside the effects of AI-supported tools. Furthermore, it remains underexplored whether these tools can produce questions with cognitive diversity comparable to that of teachers in complex and challenging science topics. In this context, it can be stated that a qualitative comparison of the questions produced by science teachers and ChatGPT on the topic of 'matter cycles' may contribute to a more informed and effective use of both teacher

practices and technological tools in education. This study examined the cognitive levels, scopes, and depth of understanding of questions generated by teachers and ChatGPT on matter cycles, allowing for the identification of their strengths and weaknesses. Furthermore, by highlighting the potential of AI-supported tools to generate questions in complex and challenging topics in science, the study provides insights for human-AI collaborative blended learning environments in education. In this regard, the research offers significant and original contributions both to science education and to digital pedagogical transformation processes. Additionally, by comparing the pedagogical skills of science teachers with the question-generation capabilities of AI-based tools, the study provides an opportunity to evaluate the impact of technological integration in education. The findings may guide teachers in improving their questioning strategies and provide insights into how AI tools such as ChatGPT can be effectively utilized in classroom learning processes. Moreover, the study is important for illuminating the design of quality questions that support students' conceptual learning in fundamental science topics, such as matter cycles. We expect that our findings will particularly contribute to the growing literature on effective classroom discourse aimed at supporting student learning through teacher questioning in science education in Turkey.

### **Science Education and ChatGPT**

The potential and impact of AI technologies in quality science education have reached a remarkable level today (Huang, 2024). With this, it has been emphasized that AI possesses the capacity to significantly transform the ways in which science is taught and learned (Cai et al., 2025). In their study review of trends and applications over the past decade, Jia et al. (2024) have demonstrated that AI has an increasingly significant impact on science education. Almasri (2024) has revealed that AI-supported tools have been extensively integrated into science education for various pedagogical purposes, including enhancing learning environments, preparing examinations, evaluating students' work, and predicting academic performance. One of the most widely used and effective applications in this domain, ChatGPT, provides information that can influence the science education of students at all levels (Cooper, 2023). Today, ChatGPT presents a mechanized human-like model in science education, providing scientific knowledge, problem-solving support, planning, and the generation of original ideas (Zhai, 2023). Due to its highly efficient ability to "understand" natural language, ChatGPT is described as a large-scale data model that closely approximates human intelligence (Binz & Schulz, 2023). In this context, the number of studies on ChatGPT in the field of science education has been steadily increasing (Ng et al., 2024; Park & Martin, 2024; Valeri et al., 2025). Alabidi et al. (2023) have highlighted that ChatGPT's personalized feedback and adaptive assessment features facilitate students' active engagement in original scientific research, the use of critical thinking skills, and participation in problem-solving activities. The study conducted by Lee and Zhai (2024) with pre-service science teachers has revealed that ChatGPT has a significant impact on student assessment.

### **Related Research**

Although various studies have been conducted on science teacher questioning (Chen et al., 2017; Ernst-Slavit & Pratt, 2017; Ong et al., 2016; Saka & Inaltekin, 2023; Soysal & Soysal, 2023), this topic has generally received limited attention in the literature. These studies have primarily focused on different instructional functions of teacher questions, including the quality of question types, the role of teacher questions in facilitating students' access to scientific content, promoting dialogic interaction, and supporting students' cognitive responses. For instance, Saka and Inaltekin (2023) conducted a video-recorded analysis of the question types used by a science teacher in the instruction of fifth grade "Force and Friction Measurement" and "Matter and Change" lessons in Turkey. The researchers found that approximately half of the questions posed by the teacher were partially appropriate or

inappropriate. Ernst-Slavit et al. (2017) examined the quality and quantity of questions posed by a teacher during a six-week science unit at a school in the Pacific Northwest of the United States. Their findings indicated that teacher questions play a significant role in facilitating students' access to scientific content. Chen et al. (2017) investigated the roles that teachers adopt when posing questions to support dialogic interaction and students' cognitive responses in argumentation practices over time. They found that as teachers assumed multiple roles, including distributor, moderator, coach, and participant, in structuring questioning patterns and shaping classroom interactions, students exhibited higher levels of cognitive responses.

To date, no studies in the literature have specifically investigated teachers' questions on matter cycles. Among the few studies addressing teachers in this context, Hoppe et al. (2020) investigated the ability of 77 science teachers to diagnose students' pre-existing understandings of matter cycles. The researchers found that teachers experienced difficulties in accurately interpreting students' understandings of ecological concepts. In another study, Hung et al. (2021) explored how an experienced science teacher implemented think-aloud and metavisualization strategies while teaching the carbon cycle. The researchers identified the relevant knowledge and skills involved in the teacher's metavisualization and, by focusing on the interactions among these knowledge and skills, highlighted critical points in the use of metacognitive strategies essential for effective metavisualization. Additionally, Lee et al. (2019) aimed to investigate how 67 science teachers understand complex topics, such as the water cycle, from a systems thinking perspective. The researchers found that teachers demonstrated varying levels of systems thinking, ranging from beginner to intermediate. It was observed that teachers experienced difficulties in completing systems thinking tasks, including identifying components and processes, recognizing multiple interactions and relationships among subsystems and hidden dimensions, and understanding the human impact on the water cycle system.

Research on AI and ChatGPT in science education has gained significant momentum following the pandemic. However, studies focusing specifically on questioning practices remain limited. In one of these studies, Lee and Zhai (2024) analyzed the lesson plans developed by 29 preservice teachers and evaluated how they integrated ChatGPT into science learning activities. The results indicated that the preservice teachers expected ChatGPT to provide opportunities for high-quality question posing and formative assessment. Shi et al. (2024) examined the usage patterns of six science teachers with an AI-supported inquiry-based intelligent tutoring system (Inq-ITS) and their perceptions of AI integration. The researchers found that all teachers acknowledged the pedagogical potential of AI support but encountered various challenges in its use. Furthermore, they highlighted that considering teachers' diverse perceptions of AI integration is crucial for supporting its effective incorporation into instructional practices. Abdelghani et al. (2025) examined middle school students' abilities to pose effective questions and critically evaluate ChatGPT's responses. The researchers found that students relied excessively on ChatGPT in both the questioning and response evaluation phases. Moreover, many students experienced difficulties in using questions aligned with learning objectives and in assessing the quality of responses. Küchemann et al. (2023) aimed to identify the characteristics and limitations of using ChatGPT 3.5 for developing physics tasks in the domain of kinematics among preservice teachers. In this experimental study, one group of preservice teachers created physics tasks using textbooks, while another group used ChatGPT. The results indicated no significant difference in task accuracy between the groups; however, those using textbooks achieved higher clarity and more frequently contextualized their questions in meaningful ways. Yang et al. (2025) compared three dialogue conditions in Korea in which students could either engage in question exchanges with a peer, pose questions to a GPT agent, or respond to questions

presented by an ITS agent through a chat interface. The results of this experimental study indicated that students in the peer and GPT groups, who were able to pose questions, demonstrated significantly higher post-dialogue performance in absolute scores compared to the ITS group, which only responded to questions. Moreover, the number of questions posed and the level of trust in the dialogue were found to be positively correlated with individual differences in performance.

### **Research Problem and Purpose**

Science education aims to enable students to understand nature from a systems-based perspective. In this context, the topic of “matter cycles” emerges as a fundamental content area that explains the transfer of energy and matter between living and non-living entities, while also fostering environmental literacy and scientific thinking skills (Zangori et al., 2017). The quality of questions used in teaching such concepts profoundly influences students’ levels of thinking, directly supporting the development of skills such as establishing relationships between concepts, conducting cause-and-effect analyses, and engaging in scientific reasoning (Hoppe et al., 2020). In the science curriculum, the topic of matter cycles (water, carbon, oxygen, and nitrogen cycles) provides a suitable platform for diverse question types and deep cognitive activities that maximize student understanding (Covitt et al., 2021; Lant et al., 2016). However, existing research indicates that teachers often fall short in effectively teaching these topics (Lee et al., 2019). It is known that teachers typically experience difficulties in understanding students’ thinking when teaching abstract topics such as matter cycles (Cardak, 2009; Malleus et al., 2016). On the other hand, advances in AI technologies offer new possibilities in instructional processes. In particular, large language models such as ChatGPT can support teachers and students in posing questions. Nevertheless, the cognitive level and pedagogical appropriateness of questions generated by such models have not yet been systematically evaluated. Although large language models such as ChatGPT have the potential to generate higher-order cognitive questions (e.g., analysis, synthesis), these capabilities have not been systematically examined in topic-specific contexts such as “Matter Cycles.” To date, no empirical study has compared the questions generated by teachers and AI tools in this area. Specifically, the systematic measurement and comparison of question quality produced by both teachers and ChatGPT in the context of matter cycles has not been conducted. Consequently, it remains unclear which components require deep cognitive engagement and how ChatGPT supports this depth. This gap creates a lack of knowledge regarding both pedagogical practices and the potential of AI-driven question development. In this context, comparing the quality of questions produced by teachers and ChatGPT regarding matter cycles reflects a critical research need, both to reveal the current state of teachers’ questioning skills and to evaluate the potential of AI-based tools in education. Additionally, teachers’ professional experience may serve as an important determinant of their pedagogical approaches in classroom practices. Therefore, differences in professional seniority may significantly influence teachers’ question-posing abilities. Furthermore, teachers’ educational level (bachelor’s, master’s, or doctoral degrees) can affect their pedagogical and content knowledge, thereby impacting the quality of their questions. Considering these variables allows for an understanding of the aspects in which teachers of different profiles generate questions that are similar to or different from those produced by ChatGPT. Therefore, the purpose of this study is to compare the questions on matter cycles generated by science teachers with varying levels of professional experience and educational backgrounds and those generated by ChatGPT in terms of cognitive level and depth of thinking. In this way, the pedagogical potential of ChatGPT is evaluated in relation to science teachers. Additionally, the study aims to reveal the strengths and weaknesses of the questions produced by ChatGPT

in terms of cognitive skills, compared to those posed by teachers. Within this framework, the research seeks to answer the following questions:

- Q1. What characteristics do the questions generated by science teachers and ChatGPT on the topic of “matter cycles” exhibit, and in what ways do they differ?
- Q2. How do the questions prepared by science teachers differ from those of ChatGPT in terms of the teachers’ professional experience?
- Q3. How do the questions prepared by science teachers differ from those of ChatGPT in terms of the teachers’ educational levels?

## **Method**

### **Research Design**

Within the scope of the study, the situations of science teachers and ChatGPT in posing quality questions on the topic of “matter cycles” were described and compared. Accordingly, the main purpose of the research was to reveal the similarities and differences in the cognitive level, conceptual depth, pedagogical structure, and clarity of meaning of the questions generated by two different sources of knowledge (humans and an artificial intelligence system). Considering this, the research was carried out using a comparative case study design, one of the qualitative research methods (Bartlett & Vavrus, 2017).

### **Research Group / Data Sources**

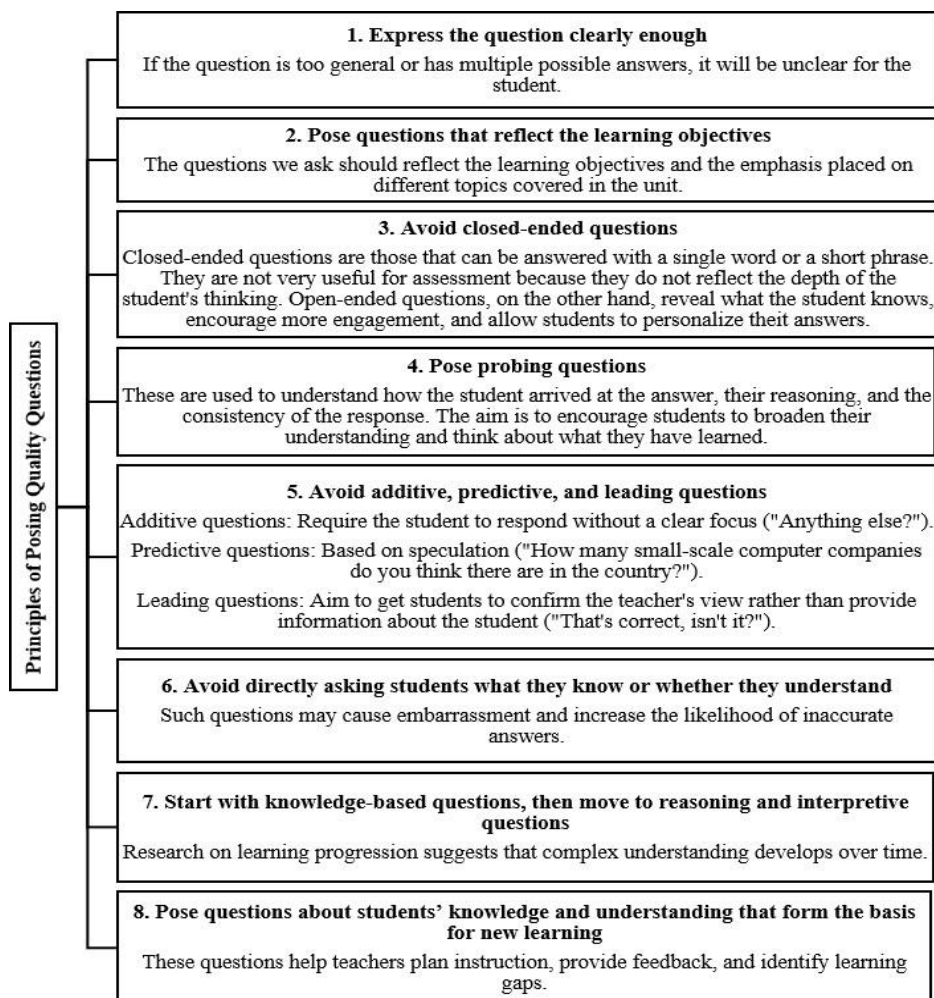
The study group consisted of 34 middle school science teachers working in Türkiye and ChatGPT (GPT-4). Of the participating teachers, 19 were female and 15 were male. In terms of professional experience, 5 teachers (14.7%) had  $\leq 3$  years, 10 teachers (29.4%) had 4–6 years, 17 teachers (50%) had 7–18 years, and 6 teachers (17.6%) had 19–30 years of seniority. Regarding educational level, 17 teachers (50%) held a bachelor’s degree, 14 (41.2%) a master’s degree, and 3 (8.8%) a doctoral degree.

### **Data Collection**

Data from the teachers were collected in written form through “The Quality Question Posing Form on Matter Cycles”, which was developed by the researchers. Teachers were asked to pose quality questions aimed at enhancing students’ understanding of the topic of “matter cycles” at the 8th-grade level. Within the scope of the study, the data were gathered during face-to-face sessions held with the participating teachers on predetermined dates. In these sessions, data collection forms prepared in line with the purpose of the study were distributed to the teachers. The participants completed the forms individually and in written form, and the completed forms were collected directly by the researchers. In this way, the data collection process was structured to allow participants to present their own statements in written form. Data from ChatGPT were obtained through texts generated in response to written prompts such as “Imagine you are a middle school science teacher teaching 8th-grade student. In this lesson, you are teaching the topic of the water cycle. What would be the best scientific questions you could pose to determine whether your students understand this topic?” Data were thus obtained as two separate sets of questions, one generated by human sources and the other by the artificial intelligence system.

### **Data Analysis**

The data were analyzed using the descriptive analysis method. In evaluating the questions generated by teachers and ChatGPT, eight criteria for quality question posing, previously developed by the researchers based on a review of the literature (Figure 1), were used.



**Figure 1.** Characteristics of Quality Questions in Classroom Assessment (McMillan, 2014)

In this analysis, the questions prepared by 34 teachers and ChatGPT for four different cycles (Water, Oxygen, Carbon, Nitrogen) were compared according to the “quality question” criteria. Data obtained from both teachers and ChatGPT were coded independently by two researchers using a double-blind procedure based on the quality question criteria. Accordingly, all data from the 34 teachers and ChatGPT were evaluated separately by the two researchers. The researchers first coded the ChatGPT data and conducted discussions until full agreement was reached on any coding discrepancies. Similarly, for the data collected from teachers, both researchers coded the data independently and then held discussions until full consensus was achieved for each teacher’s data. This procedure was systematically applied to all teacher forms. Full agreement was reached to ensure reliability. The findings were then descriptively interpreted by tabulating and graphing the frequency and percentage distributions.

## Findings

In this section, the research findings are presented regarding the characteristics of the questions generated by science teachers and ChatGPT on the topic of “matter cycles,” as well as the similarities and differences that emerged according to teachers’ professional experience and educational levels. First, the analytical findings describing the competence levels of teachers and ChatGPT in relation to the characteristics of quality questions are presented (Table 1).

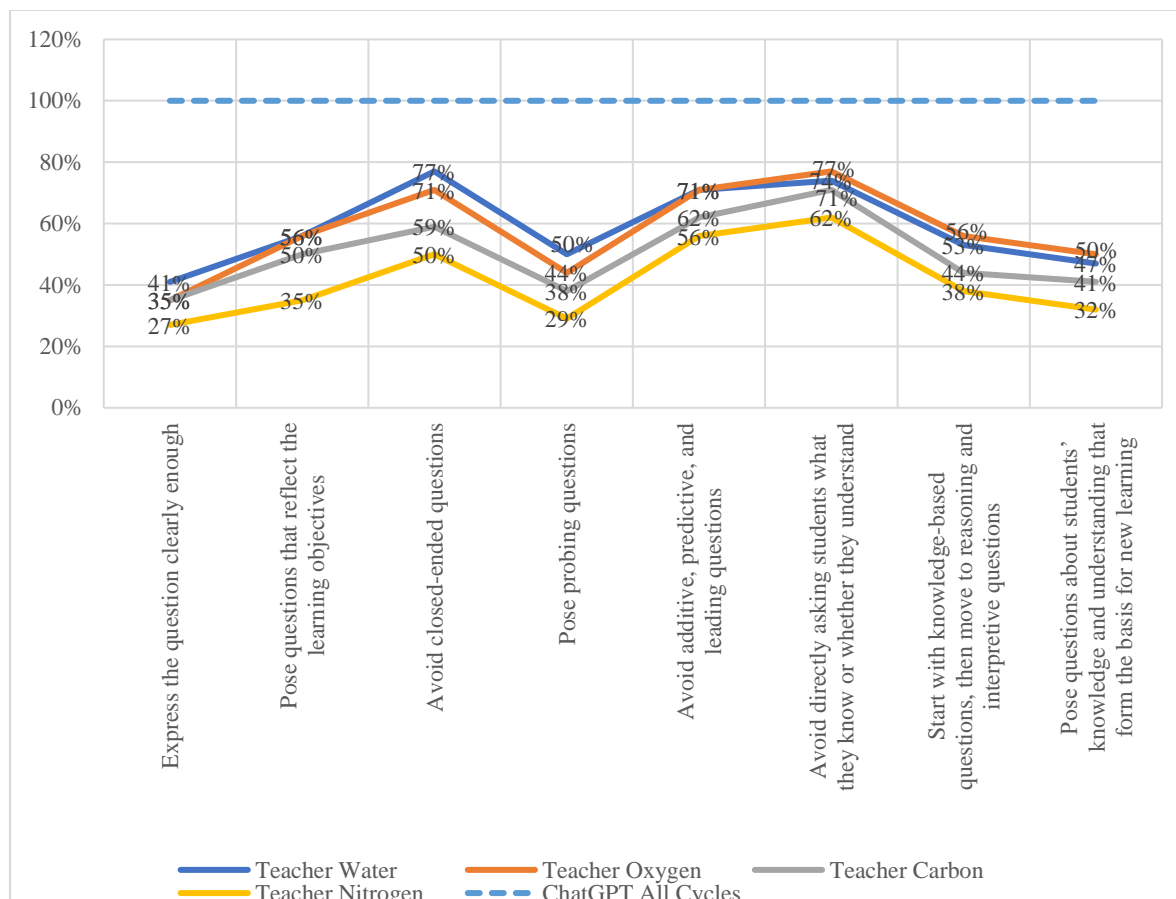
**Table 1.** General Descriptive Analysis Results of the Questions Posed by Science Teachers and ChatGPT on Matter Cycles

Source of the Question		Science Teacher (n:34)				ChatGPT			
Matter Cycles		Water	Oxygen	Carbon	Nitrogen	Water	Oxygen	Carbon	Nitrogen
Descriptive Analysis Values		n				Ability			
Principles of Posing Quality Questions	Express the question clearly enough	14	12	12	9	✓	✓	✓	✓
	Pose questions that reflect the learning objectives	19	19	17	12	✓	✓	✓	✓
	Avoid closed-ended questions	26	24	20	17	✓	✓	✓	✓
	Pose probing questions	17	15	13	10	✓	✓	✓	✓
	Avoid additive, predictive, and leading questions	24	24	21	19	✓	✓	✓	✓
	Avoid directly asking students what they know or whether they understand	25	26	24	21	✓	✓	✓	✓
	Start with knowledge-based questions, then move to reasoning and interpretive questions	18	19	15	13	✓	✓	✓	✓
	Pose questions about students' knowledge and understanding that form the basis for new learning	16	17	14	11	✓	✓	✓	✓

When Table 1 is examined, it is observed that ChatGPT met all the criteria for posing quality questions on the topic of matter cycles. It demonstrated competence in all the key features of effective questioning for each cycle, including clarity, alignment with learning objectives, open-endedness, depth, and the absence of guiding elements. In contrast, the performance of science teachers was highest in the criterion of avoiding closed-ended questions (particularly for the water cycle,  $n = 26$ , 76.5%), whereas the lowest performance was found in posing deepening questions (nitrogen cycle:  $n = 10$ , 29.4%) and eliciting students' knowledge and understanding (nitrogen cycle:  $n = 11$ , 32.4%). Overall, the nitrogen cycle appeared to be the cycle in which teachers posed the fewest quality questions. Based on these findings, it can be stated that ChatGPT demonstrates a more effective and systematic approach to generating quality questions on matter cycles, whereas teachers show individual differences in this regard, with their ability to pose effective questions decreasing particularly in abstract and more complex cycles (carbon and nitrogen).

When examining the percentage performance of teachers, the highest achievement was observed in the criterion of avoiding closed-ended questions (76.5% – water cycle), whereas the lowest performance was observed in the criteria of clearly stating the question and posing questions that form the basis for learning, as well as in the nitrogen cycle (26–32%). Additionally, the nitrogen cycle emerged as the topic in which teachers performed weakest across all criteria in generating quality questions. In contrast, ChatGPT demonstrated 100% success across all cycles and criteria.

The proportional distribution of teachers' and ChatGPT's performance in meeting the characteristics of quality questions is presented in Figure 2.



**Figure 2.** Comparison of Science Teachers' and ChatGPT's Percentages in Posing Questions According to the Criteria for Quality Questions on Matter Cycles

Figure 2 shows that ChatGPT consistently demonstrated high performance across all specified criteria. In contrast, teachers' performance varied depending on the topic and type of question. Teachers performed relatively better in the topics of water and oxygen. However, in the nitrogen topic, they clearly struggled, particularly in posing probing questions. The largest difference between teachers and ChatGPT is observed in the nitrogen cycle within the category of posing probing questions. The graph indicates that while teachers performed relatively well compared to ChatGPT in the water and oxygen topics, their performance noticeably declined in topics such as nitrogen.

Table 2 presents the descriptive analysis results of the characteristics of the questions generated by science teachers according to their professional experience.

**Table 2.** Descriptive Analysis Results of Science Teachers' Questions on Matter Cycles Based on Professional Experience

Teacher Professional Experience Codes	Matter Cycles	Principles of Posing Quality Questions							
		Express the question clearly enough	Pose questions that reflect the learning objectives	Avoid closed-ended questions	Pose probing questions	Avoid additive, predictive, and leading questions	Avoid directly asking students what they know or whether they understand	Start with knowledge-based questions, then move to reasoning and interpretive questions	Pose questions about students' knowledge and understanding that form the basis for new learning
		<i>n</i>							
≤3 year	Water	2	2	2	-	2	1	1	-
4–6 year		3	5	7	3	6	5	4	4
7–18 year		5	7	12	9	12	13	10	10
19–30 year		4	5	5	5	4	5	3	2
≤3 year	Oxygen	1	2	2	-	2	2	1	-
4–6 year		2	5	6	3	6	7	5	5
7–18 year		5	9	12	8	11	12	9	8
19–30 year		4	3	4	4	5	5	4	4
≤3 year	Carbon	-	1	2	-	2	2	1	-
4–6 year		3	4	4	3	5	6	4	3
7–18 year		7	8	10	7	10	12	8	7
19–30 year		2	4	4	3	4	4	3	4
≤3 year	Nitrogen	-	1	1	-	1	1	1	-
4–6 year		2	3	4	2	4	5	3	2
7–18 year		5	6	9	6	10	11	7	6
19–30 year		2	2	3	2	4	4	2	3

Note: ≤3 year (n:5), 4–6 year (n:10), 7–18 year (n:17), 19–30 year (n:6)

Note: Professional Career Stages ≤3 years: Commitment; 4–6 years: Career Stability; 7–18 years: Activeness; 19–30 years: Stagnation; >30 years: Withdrawal (Richter et al.'s (2011) classification of teachers' professional career stages)

Table 2 presents the descriptive analysis results of science teachers' questions on matter cycles based on professional experience. These results show that teachers with 7–18 years of experience demonstrated the highest performance in posing quality questions across all cycles and criteria (e.g., Carbon Cycle – avoiding closed-ended questions: 58.8%; Nitrogen Cycle – posing probing questions: 35.3%). In contrast, teachers with ≤3 years of experience displayed low performance in posing quality questions. Notably, none of the teachers in this group demonstrated competence in the criteria of “posing probing questions” and “posing questions about students' knowledge and understanding” across all cycles (All Cycles: 0%). Teachers with 19–30 years of experience also performed below expectations in terms of posing quality questions (e.g., Nitrogen Cycle – posing probing questions: 33.3%; Water Cycle – posing questions about students' knowledge and understanding: 33.3%). Although this group had the most experience, they fell behind the 7–18 years group in most quality question criteria. This

suggests that professional experience may plateau after a certain point. More senior teachers may also face challenges in adopting innovations. Based on these findings, teachers with mid-level experience (7–18 years) appear to be the most successful group in posing quality questions. This suggests that they may be open to contemporary pedagogical approaches. Furthermore, ChatGPT consistently demonstrated superiority across all criteria for posing quality questions (100%) compared to all teachers experience groups.

Table 3 presents the descriptive analysis results of the characteristics of the questions generated by science teachers according to their educational level.

**Table 3.** Descriptive Analysis Results of Science Teachers' Questions on Matter Cycles Based on Educational Level

Teacher Educational Level Codes	Matter Cycles	Principles of Posing Quality Questions							
		Express the question clearly enough	Pose questions that reflect the learning objectives	Avoid closed-ended questions	Pose probing questions	Avoid additive, predictive, and leading questions	Avoid directly asking students what they know or whether they understand	Start with knowledge-based questions, then move to reasoning and interpretive questions	Pose questions about students' knowledge and understanding that form the basis for new learning
Bachelor	Water	5	7	11	6	10	11	6	4
Master		7	10	12	9	11	11	10	9
Doctorate		2	2	3	2	3	3	2	3
Bachelor	Oxygen	3	6	10	5	10	11	7	6
Master		7	10	12	8	12	12	10	9
Doctorate		2	3	2	2	2	3	2	2
Bachelor	Carbon	4	6	7	3	8	10	5	4
Master		6	9	10	8	11	11	8	7
Doctorate		2	2	3	2	2	3	2	3
Bachelor	Nitrogen	2	3	6	3	7	8	4	4
Master		5	7	9	5	9	10	7	5
Doctorate		2	2	2	2	3	3	2	2

Note: Bachelor (n:17), Master(n:14), Doctorate (n:3)

Table 3 presents the descriptive analysis results of science teachers' questions on matter cycles based on educational level. The results indicate that teachers with a master's degree demonstrated high competence in meeting all quality question criteria across all cycles (e.g., Water Cycle – eliciting students' knowledge and understanding: 64.3%; Carbon Cycle – posing probing questions: 57.1%). In contrast, teachers with a bachelor's degree were successful in some areas (e.g., avoiding closed-ended questions) but showed low performance in posing probing questions (Carbon Cycle: n = 3, 17.6%) and questions aimed at understanding (Water Cycle: n = 4, 23.5%). Although the number of teachers with a doctoral degree was limited, the available data suggest that this group demonstrated consistent

competence in posing quality questions. Overall, teachers with a master's degree exhibited higher performance across all quality question criteria compared to those with a bachelor's degree. Doctoral-level teachers, despite their small number, demonstrated high competence in many quality question criteria. The performance of bachelor's degree teachers remained around 20–30% across several criteria. ChatGPT, however, outperformed all groups with a 100% success rate across all criteria. These findings suggest that teachers with a master's degree may possess more current and proficient pedagogical knowledge and methods. Teachers with a bachelor's degree, relying more on theoretical foundations, appear less capable of posing probing questions. Nevertheless, ChatGPT consistently demonstrated superior overall quality in posing questions compared to all educational levels.

### Sample Responses of Science Teachers and ChatGPT

ChatGPT

Imagine you are a middle school science teacher teaching 8th-grade students. In this lesson, you are teaching the topic of the water cycle. What would be the best scientific questions you could pose to determine whether your students understand this topic?

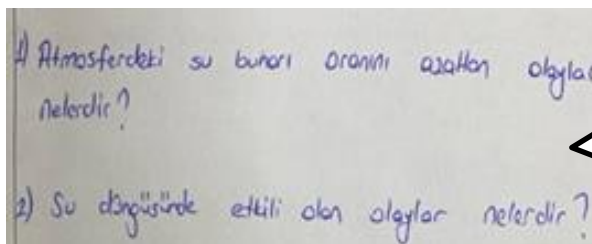
Ortaokul 8. sınıf fen bilimleri dersine giren bir öğretmen olduğunu düşün. Bu derste Su Döngüsü konusunu öğrencilerine öğretiyorsun. Öğrencilerin bu konuyu anlayıp anlamadıklarını ortaya çıkarabilmek için soracağın en iyi bilimsel sorular neler olur?

(Note: The prompt entered into ChatGPT is shown above.)

**Table 4.** Sample Responses of Science Teachers and ChatGPT

#### Matter Cycles

##### Sample Teacher Responses



- 1) What events reduce the amount of water vapor in the atmosphere?
- 2) What events play a role in the water cycle?

##### Sample ChatGPT Responses

#### Water

##### A. Temel Kavram ve Süreçleri Anlama (Bilgi – Kavrama Düzeyi)

1. "Su döngüsünde buharlaşma, yoğunlaşma ve yağış süreçleri nasıl sıralanır? Her birinin ne işe yaradığını açıklayınız."

Amaç: Öğrencinin döngüdeki temel süreçleri sıralayıp işlevini anlaması.

2. "Su döngüsünün gerçekleşmesi için Güneş neden gereklidir?"

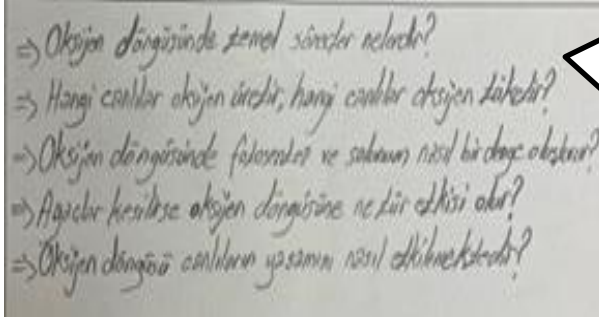
Amaç: Enerji kaynağının rolünü fark ettirme.

3. "Bitkiler su döngüsüne hangi yolla katkı sağlar?"

Beklenen yanıt: Terleme (transpirasyon).

- A. Understanding Fundamental Concepts and Processes (Knowledge – Comprehension Level)
1. How are the processes of evaporation, condensation, and precipitation sequenced in the water cycle? Explain the role of each.
  2. Why is the Sun necessary for the water cycle to occur?
  3. How do plants contribute to the water cycle?

### Sample Teacher Responses



- What are the main processes of the oxygen cycle?
- Which organisms produce oxygen, and which consume it?
- How do photosynthesis and respiration create a balance in the oxygen cycle?
- What effect would tree cutting have on the oxygen cycle?
- How does the oxygen cycle affect the survival of living organisms?

### Oxygen Sample ChatGPT Responses

#### C. İnsan Etkisi ve Ekosistemsel Düşünmeyi Ölçen Sorular

7. "Fabrika bacalarından çıkan dumanlar oksijen döngüsünü nasıl etkileyebilir?"

🔗 Amaç: Kirlenmelerin fotosentez yapan canlılara etkisi yoluyla oksijen üretimini azaltabileceğini göstermek.

8. "Oksijen döngüsü bozulursa canlı yaşamı nasıl etkilenir?"

🔗 Beklenen: Solunum zorlaşır, yaşam süreleri kısalır, ekosistemler zarar görür.

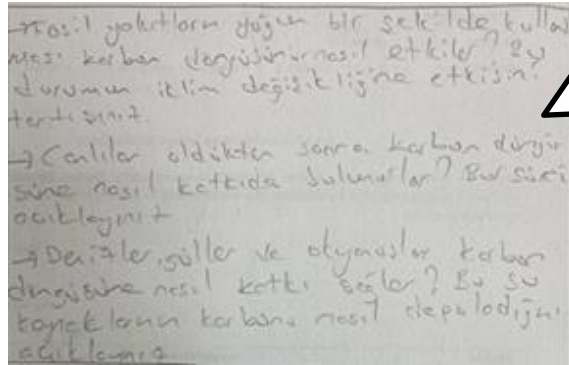
9. "Ayrıştırıcı canlıların oksijen döngüsüne katkısı nedir?"

🔗 Yanıt: Organik maddeleri parçalarken oksijen tüketirler; döngüdeki tüketici gruplardır.

#### C. Questions Measuring Human Impact and Ecosystem Thinking

7. How can smoke from factory chimneys affect the oxygen cycle?  
8. How would the disruption of the oxygen cycle affect living organisms?  
9. What is the contribution of decomposer organisms to the oxygen cycle?

### Sample Teacher Responses



- How does the intensive use of fossil fuels affect the carbon cycle?
- How do organisms contribute to the carbon cycle after they die?
- How do seas, lakes, and oceans contribute to the carbon cycle?

### Sample ChatGPT Responses

### Carbon

#### D. Yüksek Düzeyli Düşünme ve Bilimsel Yorumlama (Sentez – Değerlendirme Düzeyi)

12. "Tüm canlıların karbon döngüsündeki ortak rolü nedir? Açıklayınız."

🔗 Yanıt: Yaşarken beslenme ve solunumla karbon alıp verirler, öldüklerinde ayrışarak karbon katkısı sağlarlar.

13. "Karbon döngüsü bozulursa oksijen döngüsü etkilenir mi? Neden?"

🔗 Amaç: Döngüler arası ilişkiyi sorgulamak.

14. "Dünyada karbon miktarı neredeyse sabit olmasına rağmen atmosferdeki karbon oranı nasıl artabilir?"

🔗 Yanıt: Karbonun bazı havuzlarda (fosil yakıtlar) birikmiş olması ve yakıldığında atmosfere geçmesi.

15. "Karbon döngüsünün sağlıklı işlemesi için insanlar ne gibi önlemler alabilir?"

🔗 Beklenen: Ağaçlandırma, yenilenebilir enerji, fosil yakıt kullanımını azaltma.

#### D. Higher-Order Thinking and Scientific Interpretation (Synthesis – Evaluation Level)

12. What is the common role of all living organisms in the carbon cycle? Explain.  
13. If the carbon cycle is disrupted, would the oxygen cycle be affected? Why?  
14. Although the total amount of carbon on Earth is nearly constant, why might the carbon levels in the atmosphere be increasing?  
15. What measures can humans take to ensure the healthy functioning of the carbon cycle?

### Sample Teacher Responses

1. Toprakta ki azotun bitkiler tarafından kullanılır hale gelmesini sağlayan canlılar nelerdir ve hangi görevleri yapar?  
2. Tarımda sürekli kimyasal gübre kullanmak bu canlıları nasıl etkiler?

- What are the living things that make the nitrogen in the soil available for plants, and what tasks do they perform?
- How does the continuous use of chemical fertilizers in agriculture affect these living beings?

### Sample ChatGPT Responses

#### Nitrogen

##### D. Sentez ve Bilimsel Yorumlama (Üst Düzey Düşünme)

10. "Tüm canlılar için gerekli olan azotun atmosferde bol olmasına rağmen canlılar bu azotu nasıl kullanmak için dolaylı yollara ihtiyaç duyar?"  
Amaç: Fiziksel varlık ile biyolojik kullanılabilirlik arasındaki farkı kavramak.
11. "Azot döngüsünün durması halinde ekosistemlerde ilk hangi canlı grupları etkilenir? Neden?"  
Beklenen: Bitkiler → azot alamaz → üretici basamak çöker → tüm besin zinciri etkilenir.
12. "Azot döngüsünün sağlıklı işleyişi ile karbon döngüsü arasında nasıl bir ilişki olabilir?"  
Amaç: Döngülerin birbirine bağımlı olduğunu kavramak.

##### D. Synthesis and Scientific Interpretation (Higher-Order Thinking)

10. Although nitrogen is abundant in the atmosphere, why do living organisms need indirect pathways to use this essential element?
11. If the nitrogen cycle were to stop, which groups of organisms in ecosystems would be affected first, and why?
12. What relationship might exist between the healthy functioning of the nitrogen cycle and the carbon cycle?

### Discussion

This study revealed that in the context of matter cycles, a topic in the science education curriculum where students often experienced learning difficulties, science teachers and ChatGPT exhibited notable differences in posing quality questions.

As a prominent finding, ChatGPT demonstrated a systematic performance that met the criteria of posing quality questions across all matter cycles. In contrast, individual differences were observed in teachers' performances, and their ability to pose effective questions was relatively low, particularly in more abstract and complex cycles such as the carbon and nitrogen cycles. The lower performance of teachers in posing probing questions aimed at examining students' knowledge and understanding, especially in these more complex cycles (carbon and nitrogen), reflected the challenges encountered in teaching these topics. This situation necessitated a discussion on both the need to enhance teachers' pedagogical knowledge and strategies and the potential of AI-based tools to play a complementary role in the teaching process. The research findings were consistent with studies in the literature that had revealed teachers' limitations in posing quality questions, particularly in complex and abstract scientific processes (Eshach et al., 2014; Zhang et al., 2010). In this study as well, the low level of performance on the criteria of probing and questioning student knowledge, particularly in the nitrogen cycle, aligned with these findings. In contrast, the fact that ChatGPT met all the established criteria for quality questions demonstrated that AI tools possessed a systematic and standardized capacity for generating questions. This result emphasized the potential of AI-based tools to support the questioning processes in science education. Teachers' relatively high performance in the criterion of avoiding closed-ended questions could be considered a positive finding in terms of orienting toward inquiry-based approaches in the teaching process. However, their low performance in questions that engaged higher-order cognitive processes, such as posing probing questions and eliciting students' knowledge and understanding, could be related to their level of pedagogical content knowledge. This indicated that teachers required more comprehensive support in professional development programs focused on enhancing their ability to pose quality questions.

One notable finding of this study was that teachers exhibited significant differences in their ability to pose quality questions based on their professional experience. Analyses revealed

that teachers with 7–18 years of experience achieved the highest performance in effectively posing questions on matter cycles. This group demonstrated relatively higher competence across all cycles and all quality question criteria, indicating that the balance between professional experience and pedagogical flexibility positively contributed to the production of quality questions. This could be explained by the fact that mid-career teachers possessed established professional competencies while remaining open to innovative pedagogical approaches. In contrast, teachers with  $\leq 3$  years of experience showed very low competence in effective question posing, and notably demonstrated no performance in the criteria of posing probing questions and questions about students' knowledge and understanding. This finding suggested that in the early years of their careers, teachers might not have achieved sufficient integration between their subject matter knowledge and pedagogical content knowledge, making it difficult for them to generate questions that engaged higher-order cognitive processes. Studies on the professional challenges of novice teachers supported this result (Zhukova, 2018). The literature also emphasized that beginning teachers often relied on routine, surface-level, and knowledge-focused questions, and were generally inadequate in producing probing questions that supported inquiry-based learning (Belen & Caballes, 2020; Boakye & Ampiah, 2017; Davis et al., 2006). It was noteworthy that the performance of the most experienced group, teachers with 19–30 years of experience, lagged behind that of mid-career teachers. This indicated that experience in the teaching profession did not always increase linearly, and that pedagogical development might plateau after a certain point, or that resistance to adopting innovative approaches could occur. Indeed, some studies indicated that teachers with long-term professional experience might face challenges in adapting to innovations and tended to maintain traditional teaching approaches (Parsons et al., 2018). This further emphasized the importance of continuously involving teachers in professional development processes. Despite all these differences, it was noteworthy that ChatGPT consistently demonstrated superiority across all specified criteria. The AI tool's capacity to pose systematic, clear, and goal-oriented questions eliminated individual variations arising from human factors and provided a standardized quality of "good questions" (Scaria et al., 2024). In this context, it could be suggested that ChatGPT might serve as a supportive tool for teachers, particularly in helping novice teachers enhance their skills in posing quality questions.

Another finding of the study revealed that teachers' proficiency in posing quality questions about matter cycles varied according to their educational level. It was observed that teachers with a master's degree demonstrated high performance across all cycles and most quality question criteria. This finding suggested that teachers with a master's education were more current and competent in terms of pedagogical content knowledge and instruction. In contrast, teachers with a bachelor's degree showed adequate performance in some areas, particularly in the criterion of avoiding closed-ended questions, but exhibited limited competence in posing probing and understanding-oriented questions. This indicated that bachelor's degree teachers possessed pedagogical foundations based on theoretical knowledge but might have faced difficulties in generating questions that engage students' higher-order cognitive processes. Wong and Luft (2015) supported this expectation, highlighting that the instructional behaviors of novice teachers were part of their prior experiences. Therefore, this suggested that these bachelor's degree novice teachers were inexperienced in asking questions that reveal students' thinking processes, as posing questions that deepen student knowledge requires experience. Many newly graduated teachers with a bachelor's degree experienced self-confidence issues due to fear of losing classroom authority when posing thought-provoking questions that stimulate discussion (Kwok, 2016). Although the number of teachers with a doctoral degree was limited and the data were therefore restricted, the available evidence indicated that this group demonstrated consistent performance in posing quality questions. This suggested that

the integration of teaching experience and advanced academic education supported pedagogical knowledge coherence. Moreover, ChatGPT was able to produce consistent and high-quality questions compared to teachers at all educational levels. The AI tool's systematic and quality-focused questioning performance minimized differences arising from individual educational levels and provided a standardized "good question" quality in the teaching process. These results indicated that AI-based tools could provide complementary support in lesson preparation processes, particularly for teachers with limited pedagogical experience or gaps in certain areas.

### **Conclusion**

This study revealed that ChatGPT was capable of systematically and consistently generating quality questions on matter cycles, whereas teachers demonstrated a more limited performance, particularly in abstract and complex cycles. ChatGPT proved to be especially strong in areas where teachers tended to experience difficulties, such as probing, interpretation, and questions based on prior knowledge. Another result was that teachers with mid-level professional experience demonstrated the highest performance in posing quality questions, which clearly showed the importance of the interaction between professional experience and openness to pedagogical innovations. In addition, teachers with higher educational levels were found to be more competent in terms of pedagogical knowledge and methods, whereas teachers with only a bachelor's degree appeared to need support in generating higher-order cognitive questions. Nevertheless, the consistency provided by ChatGPT was considered valuable as a tool for balancing pedagogical differences in instructional processes.

### **Limitations**

The inclusion of 34 teachers in the study may be considered limited for representing the entire population of teachers. Therefore, this situation should be supported by future studies with larger and more diverse participant groups. ChatGPT was asked to generate questions as if thinking like a teacher. Since AI does not possess actual teaching experience or pedagogical judgment, the questions produced are solely based on the model's training data. Both teachers and ChatGPT generated questions in written form. This method may not reflect the variety of questions teachers could produce orally or in dynamic classroom situations, nor their spontaneous pedagogical decisions. Teachers' performance could also be affected by factors such as motivation, attention, and time constraints while preparing questions. In contrast, ChatGPT is not influenced by human factors such as time or fatigue. Moreover, if only one version of ChatGPT was used, results may vary with different versions or other AI models; this limitation affects the generalizability of the findings.

### **Implications**

In line with these results, it is recommended to implement systematic in-service training and pedagogical guidance programs to strengthen teachers' skills in posing quality questions on matter cycles. Increasing teachers' capacity to develop effective questions, particularly on more abstract and complex topics such as the carbon and nitrogen cycles, is crucial for supporting students' higher-order cognitive processes. Furthermore, integrating AI-based tools such as ChatGPT into lesson preparation in a way that supports teachers may contribute to enhancing pedagogical strategies and minimizing individual differences during instruction. In this context, the combined use of pedagogical training and technological tools is expected to both enhance teachers' competencies in generating quality questions and enrich students' learning experiences. Based on the results, it is also recommended to develop in-service training programs targeting inexperienced teachers to achieve quality question-posing objectives for complex and challenging topics such as matter cycles, as well as programs that support more experienced teachers in adapting to pedagogical innovations. Additionally, in-

service training programs aimed at strengthening the ability of teachers with only a bachelor's degree to probe and assess students' understanding are suggested. Finally, it is recommended that AI tools such as ChatGPT be used in lesson preparation to support these teacher groups.

**Ethics Approval:** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki.

**Conflict of Interest:** The authors declare no competing interests.

**Financial Support:** No financial support was received for this research.

## References

- Abdelghani, R., Murayama, K., Kidd, C., Sauzéon, H., & Oudeyer, P. Y. (2025). Investigating middle school students question-asking and answer-evaluation skills when using chatgpt for science investigation. *arXiv preprint arXiv:2505.01106*.
- Adel, A., Ahsan, A., & Davison, C. (2024). ChatGPT promises and challenges in education: Computational and ethical perspectives. *Education Sciences*, 14(8), 814-841. <https://doi.org/10.3390/educsci14080814>
- Alabidi, S., AlArabi, K. S., Alsalhi, N. R., & al Mansoori, M. (2023). The dawn of ChatGPT: Transformation in science assessment. *Eurasian Journal of Educational Research*, 106(106), 321-337. <https://doi.org/10.29333/ejedus/11400>
- Almasri, F. (2024). Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research. *Research in Science Education*, 54(5), 977–997. <https://doi.org/10.1007/s11165-024-10176-3>
- Barrutia, O., Ruíz-González, A., Villarroel, J. D., & Díez, J. R. (2021). Primary and secondary students' understanding of the rainfall phenomenon and related water systems: A comparative study of two methodological approaches. *Research in Science Education*, 51(Suppl 2), 823-844. <https://doi.org/10.1007/s11165-019-9831-2>
- Bartlett, L., & Vavrus, F. (2017). Comparative case studies: An innovative approach. *Nordic Journal of Comparative and International Education*, 1(1), 5–17. <https://doi.org/10.7577/njcie.1929>
- Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education*. Horizon Research, Inc.
- Belen, J. L., & Caballes, D. G. (2020). The beginning teachers 'challenges in an inquiry-based approach to teaching science: provision for a special science research elective course. *CiiT International Journal of Software Engineering and Technology*, 12, 38-44.
- Bielik, T., & Yarden, A. (2016). Promoting the asking of research questions in a high-school biotechnology inquiry-oriented program. *International Journal of STEM Education*, 3(1), 15. <https://doi.org/10.1186/s40594-016-0048-x>
- Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences of the United States of America*, 120(6). <https://doi.org/10.1073/pnas.2218523120>
- Boakye, C., & Ampiah, J. G. (2017). Challenges and solutions: The experiences of newly qualified science teachers. *Sage Open*, 7(2), 1-10. <https://doi.org/10.1177/2158244017706710>
- Cai, Y., Deng, Q., Lv, T., Zhang, W., & Zhou, Y. (2025). Impact of GPT on the academic ecosystem. *Science & Education*, 34(2), 913-931. <https://doi.org/10.1007/s11191-024-00561-9>
- Cardak, O. (2009). Science students' misconceptions of the water cycle according to their drawings. *Journal of Applied Sciences*, 9(5), 865-873.
- Chin, C. (2007). Teacher questioning in science classrooms: Approaches that stimulate productive thinking. *Journal of Research in Science Teaching*, 44(6), 815-843.

- <https://doi.org/10.1002/tea.20171>
- Chen, Y. C., Hand, B., & Norton-Meier, L. (2017). Teacher roles of questioning in early elementary science classrooms: A framework promoting student cognitive complexities in argumentation. *Research in Science Education*, 47(2), 373-405. <https://doi.org/10.1007/s11165-015-9506-6>
- 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>
- Covitt, B. A., Parker, J. M., Kohn, C., Lee, M., Lin, Q., & Anderson, C. W. (2020). Understanding and responding to challenges students face when engaging in carbon cycle pool-and-flux reasoning. *The Journal of Environmental Education*, 52(2), 98-117. <https://doi.org/10.1080/00958964.2020.1847882>
- Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research*, 76(4), 607-651. <https://doi.org/10.3102/00346543076004607>
- Düsing, K., Asshoff, R., & Hammann, M. (2019). Students' conceptions of the carbon cycle: Identifying and interrelating components of the carbon cycle and tracing carbon atoms across the levels of biological organisation. *Journal of Biological Education*, 53(1), 110-125. <https://doi.org/10.1080/00219266.2018.1447002>
- Eliasson, N., Karlsson, K. G., & Sørensen, H. (2017). The role of questions in the science classroom—how girls and boys respond to teachers' questions. *International Journal of Science Education*, 39(4), 433-452. <https://doi.org/10.1080/09500693.2017.1289420>
- Eshach, H., Dor-Ziderman, Y., & Yefroimsky, Y. (2014). Question asking in the science classroom: Teacher attitudes and practices. *Journal of Science Education and Technology*, 23(1), 67-81. <https://doi.org/10.1007/s10956-013-9451-y>
- Forster, C., & Penny, J. (2020). Questioning with student teachers. *Science Teacher Education*, 88(12), 5-9.
- Ghafar, Z. N., & Hazaymeh, O. (2024). The effective questioning in the classroom: An overview of the techniques used by instructors. *International Journal of Childhood Education*, 5(2), 1-14. <https://doi.org/10.33422/ijce.v5i2.676>
- Hoppe, T., Renkl, A., Seidel, T., Rettig, S., & Riess, W. (2020). Exploring how teachers diagnose student conceptions about the cycle of matter. *Sustainability*, 12(10), 4184. <https://doi.org/10.3390/su12104184>
- Huang, X. (2024). Application of artificial intelligence APP in quality evaluation of primary school science education. *Educational Studies*, 50(6), 1215-1235. <https://doi.org/10.1080/03055698.2022.2066462>
- Hung, J.-Y., Chang, H.-Y., & Hung, J.-F. (2021). An experienced science teacher's metavisualization in the case of the complex system of carbon cycling. *Research in Science Education*, 51, 493-521. <https://doi.org/10.1007/s11165-018-9804-x>
- Idika, M. I. & Nsifor, G. J.(2025). Use of digitalisation in science education: a case study of STEM teachers in the Ibadan metropolis, Kenya. *International Journal of Education, Technology and Science*, 5(2), 209-223. <https://doi.org/10.5281/zenodo.15499135>
- Jia, F., Sun, D., & Looi, C. K. (2024). Artificial intelligence in science education (2013–2023): Research trends in ten years. *Journal of Science Education and Technology*, 33(1), 94-117. <https://doi.org/10.1007/s10956-023-10077-6>
- Kayima, F., & Jakobsen, A. (2020). Exploring the situational adequacy of teacher questions in science classrooms. *Research in Science Education*, 50(2), 437-467. <https://doi.org/10.1007/s11165-018-9696-9>
- Kuchemann, € S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). Physics task development of prospective physics teachers using ChatGPT. *arXiv preprint arXiv:2304.10014*.

- Kwok, A. H. (2016). *Exploring beginning teachers' beliefs, actions, and influences on classroom management* (Unpublished Doctoral dissertation). University of Michigan.
- Lant, C., Pérez-Lapeña, B., Xiong, W., Kraft, S., Kowalchuk, R., & Blair, M. (2016). Environmental systems simulations for carbon, energy, nitrogen, water, and watersheds: Design principles and pilot testing. *Journal of Geoscience Education*, 64(2), 115-124. <https://doi.org/10.5408/14-004.1>
- Lee, T. D., Gail Jones, M., & Chesnutt, K. (2019). Teaching systems thinking in the context of the water cycle. *Research in Science Education*, 49(1), 137–172. <https://doi.org/10.1007/s11165-017-9613-7>
- Lee, G. G., & Zhai, X. (2024). Using ChatGPT for science learning: A study on pre-service teachers' lesson planning. *IEEE Transactions on Learning Technologies*, 17, 1683-1700. <https://doi.org/10.1109/TLT.2024.3401457>
- Malleus, E., Kikas, E., & Kruus, S. (2016). Students' understanding of cloud and rainbow formation and teachers' awareness of students' performance. *International Journal of Science Education*, 38(6), 993-1011. <https://doi.org/10.1080/09500693.2016.1175683>
- Morris, J., & Chi, M. T. (2020). Improving teacher questioning in science using ICAP theory. *The Journal of Educational Research*, 113(1), 1-12. <https://doi.org/10.1080/00220671.2019.1709401>
- Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55(4), 1328-1353. <https://doi.org/10.1111/bjet.13454>
- Nielsen, L. S. (2009). *The relationship between pedagogical content knowledge and mathematics teacher questioning strategies*. (Unpublished doctoral dissertation). Louisiana Tech University.
- Oliveira, A. W. (2010). Improving teacher questioning in science inquiry discussions through professional development. *Journal of Research in Science Teaching*, 47(4), 422-453. <https://doi.org/10.1002/tea.20345>
- Park, H. K., & Martin, S. N. (2024). Exploring the role of ChatGPT in science education for Asia-Pacific and beyond: a systematic review. *Asia-Pacific Science Education*, 10(2), 233-264. <https://doi.org/10.1163/23641177-bja10079>
- Park, J., Teo, T. W., Teo, A., Chang, J., Huang, J. S., & Koo, S. (2023). Integrating artificial intelligence into science lessons: Teachers' experiences and views. *International Journal of STEM Education*, 10(1), 61. <https://doi.org/10.1186/s40594-023-00454-3>
- Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., ... & Allen, M. (2018). Teachers' instructional adaptations: A research synthesis. *Review of Educational Research*, 88(2), 205-242. <https://doi.org/10.3102/0034654317743198>
- Rejeb, A., Rejeb, K., Appolloni, A., Treiblmaier, H., & Iranmanesh, M. (2024). Exploring the impact of ChatGPT on education: A web mining and machine learning approach. *The International Journal of Management Education*, 22(1), 100932. <https://doi.org/10.1016/j.ijme.2024.100932>
- Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, 44(1), 57-84. <https://doi.org/10.1002/tea.20163>
- Sadler, T. D., Nguyen, H., & Lankford, D. (2017). Water systems understandings: A framework for designing instruction and considering what learners know about water. *Wiley Interdisciplinary Reviews: Water*, 4(1), e1178. <https://doi.org/10.1002/wat2.1178>
- Scaria, N., Chenna, S. D., & Subramani, D. (2024). Automated educational question

- generation at different Bloom's skill levels using large language models: Strategies and evaluation. In A. M. Olney, I. A. Chounta, Z. Liu, O. C. Santos, & I. I. Bittencourt (Eds.), *Artificial intelligence in education. AIED 2024. Lecture notes in computer science* (Vol. 14830, pp. 153–167). Springer. [https://doi.org/10.1007/978-3-031-64299-9\\_12](https://doi.org/10.1007/978-3-031-64299-9_12)
- Shi, L., Ding, A.-C., & Choi, I. (2024). Investigating teachers' use of an AI-enabled system and their perceptions of AI integration in science classrooms: A case study. *Education Sciences*, 14(11), 1187. <https://doi.org/10.3390/educsci14111187>
- Soysal, Y., & Soysal, S. (2023). The art of asking good questions in the classroom: a phenomenographic study of teacher educators' recommendations. *ECNU Review of Education*, 20965311231210008. <https://doi.org/10.1177/20965311231210008>
- Talbot, R. M., Wylie, R., Dutilly, E., & Nielsen, R. (2018). The relationship between format and cognitive depth of science teacher-generated questions. *Research in the Schools*, 25(1), 35-46.
- Ugulu, I., Yorek, N., & Baslar, S. (2015). The effect of recycling education on high school students' conceptual understanding about ecology: A study on matter cycle. *Educational Research and Reviews*, 10(16), 2207-2215.
- Vale, R. D. (2013). The value of asking questions. *Molecular Biology of The Cell*, 24(6), 680-682. <https://doi.org/10.1091/mbc.e12-09-0660>
- Valeri, F., Nilsson, P., & Cederqvist, A. M. (2025). Exploring students' experience of ChatGPT in STEM education. *Computers and Education: Artificial Intelligence*, 8, 100360. <https://doi.org/10.1016/j.caeai.2024.100360>.
- Van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM Mathematics Education*, 53, 17–27. <https://doi.org/10.1007/s11858-020-01211-4>
- Wellington, J., & Osborne, J. (2001). *Language and literacy in science education*. Open University Press.
- Wong, S. S., & Luft, J. A. (2015). Secondary science teachers' beliefs and persistence: A longitudinal mixed-methods study. *Journal of Science Teacher Education*, 26(7), 619-645. <https://doi.org/10.1007/s10972-015-9441-4>
- Yang, J. W., Choe, K., Lim, J., & Park, J. (2025). Asking questions matters: Comparing the effect of learning through dialogue with peers, ChatGPT, and intelligent tutoring system. *Interactive Learning Environments*, 1–17. Online first. <https://doi.org/10.1080/10494820.2025.2488984>
- Yip, D. Y. (2004). Questioning skills for conceptual change in science instruction. *Journal of Biological Education*, 38(2), 76-83. <https://doi.org/10.1080/00219266.2004.9655905>
- Zangori, L., Peel, A., Kinslow, A., Friedrichsen, P., & Sadler, T. D. (2017). Student development of model-based reasoning about carbon cycling and climate change in a socio-scientific issues unit. *Journal of Research in Science Teaching*, 54(10), 1249-1273. <https://doi.org/10.1002/tea.21404>
- Zhai, X. (2023). ChatGPT for next generation science learning. *XRDS: Crossroads, the ACM Magazine for Students*, 29(3), 42–46. <https://doi.org/10.1145/3589649>
- Zhai, J., & Tan, A. L. (2015). Roles of teachers in orchestrating learning in elementary science classrooms. *Research in Science Education*, 45(6), 907-926. <https://doi.org/10.1007/s11165-014-9451-9>
- Zhang, M., Lundeberg, M., McConnell, T. J., Koehler, M. J., & Eberhardt, J. (2010). Using questioning to facilitate discussion of science teaching problems in teacher professional development. *Interdisciplinary Journal of Problem-Based Learning*, 4(1), 57–82. <https://doi.org/10.7771/1541-5015.1097>
- Zhukova, O. (2018). Novice teachers' concerns, early professional experiences and

development: Implications for theory and practice. *Discourse and Communication for Sustainable Education*, 9(1), 100-114. <https://doi.org/10.2478/dcse-2018-0008>