

Evaluating the Efficacy of AI-Generated Inquiry-Based Lesson Plans in Science

Nurcan Kahraman^{1*} 
Gülbin Kiyıcı² 

¹ Bursa Uludağ University, Bursa School of Education, Bursa, Türkiye, nurcankahraman@uludag.edu.tr, ror.org/03tg3eb07

² Manisa Celal Bayar University, Manisa, Türkiye, gulbin.kiyici@gmail.com, ror.org/053f2w588

*Corresponding author

Received: 01.04.2024
Accepted: 15.03.2025
Available Online: 16.04.2025

Abstract: This study aims to investigate the effectiveness of artificial intelligence in creating inquiry-based lesson plans for science education in the context of Turkey. An exploratory design was used to achieve this goal. The subject matter of middle school science was treated as a case and ChatGPT was asked to create lesson plans by specifying the objectives and duration of the lessons. Then, the plans were evaluated by the researchers, considering the rubric of the Guidelines for Assessing Inquiry-based Lesson Plans developed by Borda-Carulla and Harmen (2012). The researchers of this study evaluated the lesson plans separately, and the internal consistency coefficient was calculated as 0.83. The results suggest that the artificial intelligence-generated lesson plans were aligned with the course objectives, incorporated real-world contexts, and promoted an inquiry-based learning environment. Despite these strengths, the results of this study also suggest that some areas need to be improved, such as providing structured feedback and fostering curiosity.

Keywords: AI-Generated, Lesson Planning, Science Education, ChatGPT, Inquiry-Based Learning

1. Introduction

Artificial intelligence (AI) has become more common and attracted much attention in education. UNESCO has published *AI Competency Framework for Students* to help educators integrate AI into education as a global standard and aims to educate students to be responsible, knowledgeable users and creators of AI. The framework underlines the importance of educating students with skills to use AI effectively. The framework investigates these necessary knowledge and skills under four main categories: Human-centred mindset; Ethics of AI; AI techniques and applications; and AI system design. This framework not only serves as a guide for integrating AI into curricula but also highlights the importance of equipping students with critical and ethical awareness, preparing them for an active role in AI-driven societies (UNESCO, 2024). Although researchers have been discussing the use of AI in education for about 30 years, its educational potential is still a black box for educators (Zawacki-Richter et al., 2019). The integration of AI in education has led to the creation of more personalized and effective learning experiences through enhanced technology platforms. The role of AI in education encompasses several key dimensions, including facilitating nuanced teacher feedback, implementing automated grading systems, promoting adaptive learning techniques, and supporting distance learning frameworks (Liang et al., 2021). Besides this, it is revolutionizing teaching approaches, enhancing learning processes, and refining assessment strategies (Johnson et al., 2017). It encompasses a wide range of applications, from adaptive learning platforms that tailor content to the needs of individual learners to analytics tools that monitor student engagement and performance (Chen, et al. 2020). Additionally, AI has the potential to transform lesson planning and curriculum development, making traditional methods, which are often time-consuming and static, more dynamic and personalized, thereby reducing the burden on educators. For example, in a recent study, Cooper (2023) examined the responses of ChatGPT, an AI tool, to questions related to science education. The answers were mostly consistent with the literature. Furthermore, in the task of generating a lesson for a science subject, the researcher suggests that the ChatGPT feature, although it needs further fine-tuning, is expected to gain wide acceptance among educators. Based on this, our study aims to explore the effectiveness of AI in creating lesson plans for science education in the context of Türkiye.

To respond to a rapidly changing global landscape driven by technological progress, to show openness to new ideas, and to develop individuals capable of integrating scientific principles into everyday life, improvements in science education and training are underway, aiming to achieve global competitiveness in Türkiye. In other words, the swift advancement of science and technology, the changing needs of individuals and communities, contemporary innovations, and developments in educational approaches have directly influenced the roles individuals need to adopt. To meet these challenges, Türkiye has adopted inquiry-based learning in its science education program (MoNE, 2018). With this transformation, individuals are expected to produce knowledge, use it effectively in practical situations, solve problems, think critically, have an entrepreneurial spirit, be determined, have communication skills, empathize, and add value to society and culture. This raises a question about the inquiry-based learning strategy: "If students are to learn the methods of science, what better way than through active participation in the inquiry process itself?" (NRC, 2000). This study thus aligns AI's potential with Türkiye's educational goals, aiming to assess the quality of AI-generated lesson plans within an inquiry-based learning context.

1.1. Inquiry-based learning

As in all learning and teaching approaches, the inquiry-based learning approach is used together with other learning approaches. No learning and teaching approach alone can be effective for a long time. Nor has this approach emerged in isolation. Inquiry-based learning emerged under the influence of many important theories and it has influenced other learning strategies, such as constructivist learning (Degenhart, 2010). To demonstrate the importance of the inquiry-based learning strategy, the question "If students were to learn the methods of science, then how better to learn than through active engagement in the process of inquiry itself?" should be answered (NRC, 2000). When explaining this strategy, it is necessary to focus on what provides more than just a dictionary definition. In this context, learning benefits concerning the inquiry-based learning strategy have been identified. These benefits include scientific process skills, the nature of scientific research-questioning, and the general research-questioning process (Barrow, 2006). At the end of the IBL process, the benefits identified by the students include investigating information from life by using information steps and problem-solving skills and developing skills and attitudes that can generalize this information (Yaşar & Duban, 2009). When students learn by conducting inquiry, they go through the following stages: undertaking the research-sorting process, exhibiting curiosity, defining current information questions, simple inquiry planning and execution, proposing preliminary explanations or hypotheses, collecting evidence from observations, using evidence-based explanations, and thinking of and explaining other explanations and communications (NRC, 2000). When this process is examined carefully, it can be seen that students follow their methods and practices by putting themselves in the place of scientists who collect information about the world to accumulate information (Keselman, 2003). The provision of benefits is associated with the level of inquiry. Furthermore, the quality of the subject to be applied and the target audience also influences the level of questioning.

The relevant literature has shown that inquiry-based learning has numerous positive effects on improving students' attitudes toward the subject under study, academic performance, scientific process skills, perception of inquiry skills, motivation, application of research skills, meaning-making, and acquisition of scientific knowledge (Suduc et al., 2015; Ünlü, 2015; Varlu & Sağır, 2019; Yerlikaya, 2019). Studies conducted in Türkiye are also in line with international literature. For instance, the positive effects on achievement, perception of inquiry-learning skills, and metacognition levels after teaching the propagation of light and sound at the fifth-grade level following the inquiry-based approach can be seen in the results from the study by Varlu and Sağır (2019). At the seventh-grade level, students' scientific process skills improved after the application of inquiry-based learning in the biology unit 'Systems in our Body' and, because of a qualitative data analysis, the students could see the skills they had developed following their inquiry-based learning (Yerlikaya, 2019). In another study, students found doing

research fun, although their teachers did not give them research assignments. The perceptions of pre-service science teachers toward scientific process skills, their attitudes toward the laboratory, and their academic achievements were positively affected by the inquiry-based learning practices in the Chemistry Laboratory-II course they taught (Ünal, 2018). In the general physics course, a guided inquiry-based set of lessons on electrical current and simple electrical circuits was made and, although it had a positive effect on academic achievement, the prospective teachers did not trust the knowledge obtained at the end of the inquiry process (Yetişir, 2016). In Güler and Şahin (2018), a group of pre-service teachers had positive views about the inquiry process. An inquiry-based teaching strategy positively improved their conceptual learning and academic self-efficacy (Kayacan & Selvi, 2017). In another study, a set of teacher guidance materials (TGMs) was developed in which the stages and principles for the implementation of the guided inquiry-based teaching approach were emphasized. In another study, a set of TGMs was developed for a guided inquiry-based teaching approach (Saka, 2018). The activities in the sixth-grade science textbook, which was prepared based on the inquiry-based learning approach of the science curriculum, were analyzed using the scale of the requirements of the inquiry-based learning model. It was determined that it did not fully meet the requirements of the criteria determined in the scale (Sarioğlu et al., 2016). However, it is essential to integrate this method effectively. Studies have shown that teachers' attitudes, self-efficacy, and the appropriateness of the activities for inquiry-based learning are crucial factors (Açıkgöz et al., 2018; Dobber et al., 2017; Sarioğlu et al., 2016; Cleovoulou & Beach, 2019). Therefore, it is important to support teachers to understand inquiry-based learning and use it to design lessons (Inoue et al., 2019; Teig et al., 2018; Villardón-Gallego, 2016; Saka, 2018). Planned, programmed, and systematic teaching experiences organized in schools create an important dimension of the education process (Can, 1998). In this context, teachers in charge and teacher candidates should have knowledge of and experience in the effectiveness of AI in creating lesson plans.

1.2. Lesson planning and AI

For a teacher, in line with the objectives of the program; it is an important professional competence to regulate content, educational situations, and testing situations in the most effective way for students (Yıldırım & Yıldırım, 2020). Lesson planning directs a teacher to the right path to achieve predetermined goals. It also helps the teacher in evaluating himself to implement techniques and strategies at a further level (Khan & Elixir, 2011). By specifying instructional objectives, teachers define their purposes in terms that are clear and understandable. Well-written instructional objectives enable teachers to plan and implement their instructional strategies. Planning is arguably the most crucial task a teacher undertakes. As new teaching scenarios emerge, planning skills will remain a vital aspect of every teacher's professional life. (Cooper, 2011). Regarding science education, a well-structured lesson plan provides students with a deep understanding of scientific concepts, develops their critical thinking and problem-solving skills, and encourages them to actively participate in the scientific process (Krajcik & Czerniak, 2014).

Given the importance of lesson planning in reaching educational objectives and the challenges teachers encounter in this process, there is a need for more innovative approaches to simplify it. One possible solution is to leverage the assistance of AI (Celik et al., 2022; Cooper, 2023; Zawacki-Richter et al., 2019). Teaching is a complex, cognitively demanding, and highly creative process; AI technologies can be utilized to support competency-based learning and to enhance teachers' creativity as they plan their lessons (Pender et al., 2022). In particular, AI applications provide basic frameworks and materials that can save teachers time while also adapting to the diverse needs of students and educators. Reflecting on activities created by ChatGPT can help broaden teachers' perspectives and inspire creativity by offering various approaches to teaching a specific topic. These suggestions can encourage teachers to think outside the box and consider new and innovative ways to engage their students in the learning process (Van den Berg & du Plessis, 2023).

Recognizing the challenges teachers face in lesson planning, it becomes evident that AI's evolving capabilities offer a promising avenue to address these difficulties and enhance educational outcomes. In other words, the integration of AI in educational settings not only enables personalized learning experiences but also supports educators by lighting their workload (Luckin et al., 2016). For instance, prior research has shown that AI can aid teachers in planning and delivering lessons (Celik et al., 2022; Zawacki-Richter et al., 2019). An example of AI's influence on educational planning is found in Cooper's (2023) study, which explored ChatGPT's role in science education pedagogy. According to his research, when asked about effective science teaching strategies, ChatGPT responded consistent with established research. Moreover, the AI tool was tasked with creating instructional units, and the resulting content proved valuable for educators. This example highlights how AI applications can offer both practical and theoretical support, particularly in essential areas like lesson planning. Hence, the purpose of this study is to investigate how AI can design inquiry-based lesson plans for middle-school science courses. For this aim, the research question below was asked:

- To what extent do AI-generated inquiry-based lesson plans meet the standards of quality in terms of design and content?

2. Method

2.1. Research design

The present research is designed as an exploratory study. Exploratory research is a critical method used when a problem or context lacks clear definition or comprehension. It serves as an initial investigation to gather more information and gain a deeper understanding of the key issues or opportunities before starting more structured research projects. The primary goal of this type of research is to identify patterns, theories, or understandings rather than reach a conclusion (Creswell, 2014). Additionally, the current study utilized a self-study methodology to achieve its objectives. This methodology emphasizes the researcher's practices and places a strong emphasis on analyzing one's own experiences to gain new perspectives and make improvements (Hamilton et al., 2009).

2.2. Lesson plans

The purpose of this study was to examine the effectiveness of AI in creating inquiry-based lesson plans for science education. To achieve this goal, we investigated the science curriculum for middle-school students in Türkiye. The topic of 'Matter and its Nature' was then discussed as the subject of the case. All the concepts for each grade are presented in Table 1.

Table 1

Concepts of 'Matter and Its Nature for Each Grade

Grade	Topic	Hour	Concepts
5	Change of State of Matter	6	Melting, freezing, boiling, condensation, evaporation, sublimation, condensation
	Distinguishing Properties of Matter	6	Melting and freezing point, boiling point
	Heat and Temperature	7	Heat, temperature, heat exchange
	Heat Affects Substances	7	Expansion, contraction
6	Particle Structure of Matter	6	Particle structure, hollow structure, moving structure
	Intensity	8	Density, density unit
	Matter and Heat	8	Thermal conductivity, thermal insulation, thermal insulation materials
	Fuels	6	Solid fuels, liquid fuels, gaseous fuels, renewable and non-renewable energy sources

Table 1 (Continued)

7	Particle Structure of Matter	6	Atom (nucleus, layer, proton, neutron, electron), property of scientific knowledge, molecule
	Pure Substances	6	Element, symbols of elements, compound, compound formulas
	Mixtures	6	Homogeneous mixture, solution (solute, solvent), heterogeneous mixture, dissolution, factors affecting dissolution rate
	Separation of Mixtures	4	Evaporation, density difference, distillation
	Household Waste and Recycling	6	Domestic solid waste materials, domestic liquid waste materials, recycling, reuse
	Periodic System	4	Group, period, classification of periodic system
	Physical and Chemical Changes	4	Physical change, chemical change
8	Chemical Reactions	3	Occurrence of chemical reactions, conservation of mass
	Acids and Bases	8	Acids, bases, pH, acid rain, solutions against acid rain
	Interaction of Matter with Heat	5	Factors related to heat and self-heat
	Chemical Industry in Türkiye	4	Imported chemical products, exported chemical products, official / private institutions that contribute to the development of the chemical industry in our country, chemistry-based professions

ChatGPT-4 was used to create the lesson plans. GPT-4 (generative pre-trained transformer) is a type of generative AI language model created by OpenAI (OpenAI, 2023). Then, after providing details on grade levels, time allocated, and learning outcomes, the program was asked to create lesson plans that were appropriate for inquiry-based learning. Specifically, the prompt "Please create lesson plans that are appropriate for inquiry-based learning, considering the provided grade levels, time allocated, and learning outcomes" was provided to GPT-4. An example of the prompt given and the lesson plans created is presented below, in Figure 1 and in Table 2, respectively.

Figure 1

Example of a Prompt Given to ChatGPT

ChatGPT ▾

Please create lesson plans that are appropriate for inquiry-based learning, considering the provided grade levels, time allocated, and learning outcomes.

Grade 5.

F.5.4.1. Change of State of Matter Recommended Duration: 6 class hours Subject / Concepts: Melting, freezing, boiling, condensation, evaporation, sublimation, condensation

F.5.4.1.1. Makes inferences based on the data obtained from the experiments that substances can change state under the influence of heat. The basic difference between evaporation and boiling is explained by stating that liquids evaporate at any temperature but boil at certain temperatures.



Bellek güncellendi

Inquiry-Based Lesson Plan: Change of State of Matter (Grade 5)

Table 2*Lesson Plan Example Generated by ChatGPT*

Learning outcome	F.5.4.2.1. Determines the melting, freezing and boiling points of pure substances as a result of his/her experiments. It is emphasized that melting, freezing and boiling points are distinguishing features.	
Subject	Distinctive Properties of Matter	
Grade	5	
Recommended Period	6 lesson hours	
Concepts	Melting point, freezing point, boiling point	
1. Lesson Hour: Introduction and Basic Concepts		Discuss the three states of matter (solid, liquid, gas) and the main reasons for transitions between these states.
	Activities	Present the basic definitions of melting, freezing and boiling points and emphasize that these concepts are the distinguishing properties of substances.
	Purpose:	To explain the basic definitions and importance of melting, freezing and boiling points to students.
2. Lesson Hour: Discovery of the Melting Point		Conduct melting point experiments using various pure substances (e.g. ice, chocolate).
	Activities	Have students record their observations and collect melting point data.
	Purpose:	Explore how melting point is a distinguishing property of substances and give students an active role in the experimentation process.
3. Lesson Hour: Investigation of Freezing Point		Organize experiments to compare the freezing points of water and saline solutions.
	Activities	Ask students to record their freezing point observations and results.
	Purpose:	To show that the freezing point is a distinguishing property of pure substances and how it can vary depending on the components of the solution.
4. Lesson Hour: Boiling Point Research		Conduct experiments to observe the boiling points of different liquids (water, salt water, alcohol).
	Activities	Provide students with thermometers to record temperature changes during boiling.
	Purpose:	To examine that the boiling point is a distinguishing characteristic of substances and the effect of atmospheric pressure on the boiling point.
5. Lesson Hour: Data Collection and Analysis		Collect the data from the experiments carried out in the previous lesson time.
	Activities	Have students analyze their data in groups and determine the melting, freezing and boiling points of each substance
	Purpose:	To give students the opportunity to analyze data and draw scientific conclusions.
6. Lesson Hour: Presentation and Evaluation		In groups, give students the opportunity to present the results of their experiments to the class.
	Activities	Discuss as a class how melting, freezing and boiling points can be used to distinguish substances.
	Purpose:	To reinforce what students have learned and to share knowledge, and to emphasize the importance of melting, freezing and boiling points.

2.3. Rubric for evaluation

The "Guidelines for Assessing Inquiry-Based Lesson Plans" (<https://www.scribd.com/document/379702927/P2P-Rubric>), developed by Borda-Carulla and Harmen (2012), were used to evaluate the lesson plans. This rubric is derived from those created by the Fibonacci Project, which ran from 2010 to 2013. The Fibonacci Project aimed to widely promote inquiry-based science and mathematics education across the European Union. The project involved developing and testing a unified approach to inquiry-based teaching and learning in science and mathematics and implementing a dissemination process with 12 reference centers and 24 twin centers throughout Europe, tailored to local contexts (Fibonacci, 2012). The project developed three different rubrics on inquiry-based learning for teachers through interviews with students and teachers, classroom observations, etc. They then provided guidelines for evaluating inquiry-based lesson plans.

The rubric is designed to assess various aspects of inquiry-based learning-including instructional practices, real-world connections, inquiry processes, collaboration, documentation, and evaluation rubric uses a five-point scale (1 = not at all, 5 = completely). An examination of the rubric items, as shown in Table 3, reveals that it comprehensively addresses all of the critical criteria needed to assess inquiry-based learning processes, making it a relevant and thorough assessment tool. The clear and explicit definitions of each criterion also ensure consistent understanding among evaluators and students, thus facilitating objective assessments. Therefore, this rubric was used to evaluate the lesson plans in the current study. While direct evidence validating this specific rubric is not available, its foundation in the well-regarded Fibonacci Project, and its comprehensive coverage of inquiry-based learning criteria, suggest that it is a suitable tool for evaluating inquiry-based lesson plans in this context.

Ideally, an evaluation should yield consistent results regardless of who performs the scoring, and the outcomes should be similar regardless of the time and location of the assessment. However, achieving this level of consistency is challenging. The greater the consistency of scores across different raters and occasions, the more reliable the assessment is considered to be (Jonsson & Svingby, 2007). The researchers of this study evaluated the lesson plans separately and the internal consistency coefficient, measured using Cohen's Kappa, was calculated as .83. The items of the rubric can be seen in Table 3. The Kappa statistic was developed to determine the degree of agreement between two raters scoring at the classification level (Cohen, 1988). The Kappa statistic assumes these categorized objects or individuals are independent, that raters' scores are independent of each other, and that the scoring categories are independent of each other (Brennen & Prediger, 1981).

3. Results

The lesson plans created by the AI were first analyzed with a focus on content. This involved a thorough review of the training material by two experts to assess its relevance, accuracy, and comprehensiveness. The aim was to ensure that the plans created by the AI accurately addressed the learning outcomes. According to the findings, all the content of the lesson plans was aligned with the relevant objectives. In addition, the lesson plans demonstrated a clear understanding of the subject matter and provided a structured approach to achieve the intended learning outcomes.

To evaluate the lesson plans, the researchers used the Guidelines for Assessing Inquiry-based Lesson Plans, developed by Borda-Carulla and Harmen (2012). The scoring matrix uses a 5-point scale and has six dimensions. These were general, real-world connections, inquiry, collaboration, documentation, and evaluation. The mean scores of the lesson plans for each criterion are presented in Table 3.

Table 3*The Mean Scores of the Lesson Plans*

Criteria	Mean
General	
The lesson plan is definitely an inquiry-based lesson and/or utilizes at least 1 of the methods presented in this MOOC (e.g. writing to learn activities, the Notebook of Experience, Philosophy for Children, the Experimenter Report – or a more general IBSE activity).	3.52
The lesson plan is clear, easy to understand and concise in its manner.	4.05
The lesson plan is well aligned with its learning outcomes: activities and assessment clearly link with the defined learning outcomes and allow the teacher to determine by the end of the lesson(s) if the objectives have been achieved.	4.81
The lesson plan includes a good mix between inquiry, scaffolding for learning and hands-on tasks where appropriate. There is a good mix of activities used and none of the activities take up more than 35% of the time (see the Learning Designer pie chart for this).	3.86
The learning design shows that students are offered regular, structured and authentic feedback to improve the quality of their work.	2.86
Average	3.82
Real-world connections	
The starting point of the lesson stimulates curiosity. For example, it is an activity for a science lesson that relates to an expedition in the wild.	3.00
The learning is embedded in real world contexts (the task is connected to students' lives) and the activities allow students to understand how the procedures they learn are applied in real-life situations. For example, a chemistry activity that utilizes everyday materials, which can be found at home.	4.14
Average	3.57
Inquiry	
The teacher facilitates students' development by allowing them to explore and develop their own ideas and ask questions about the task at hand.	3.86
Students are encouraged to carry out an investigation. For example, they are actively involved in collecting information (either from real objects or from secondary sources such as books, posters, websites).	4.52
Students are encouraged to draw conclusions, check their results, compare results with predictions and document their progress.	4.48
Average	4.29
Collaboration	
The instruction time is well balanced between whole group and small group instruction, including a good mixture between individual and team work.	3.62
The lesson includes group discussions, in which pupils can freely express their thoughts and ask questions.	4.57
Average	4.1
Documentation	
The lesson makes sure that observation and data are recorded and collected in a systematic way.	4.14
Students are encouraged to present their findings to the class.	4.62
Average	4.38
Evaluation	
The lesson plan clearly indicates that the class has a chance to evaluate the inquiry activity together.	3.62

As can be seen in Table 3, the overall mean score of the evaluation was 3.82. Of the six dimensions, real-world connections had the lowest mean score of 3.57, and documentation had the highest mean score of 4.38. In terms of items, the highest score belonged to 'The lesson plan is well aligned with its learning outcomes: activities and assessment are linked to the defined learning outcomes and allow the teacher to determine at the end of the lesson(s) whether the objectives have been achieved' with 4.81. Encouraging students to present the results of the lesson followed with an average of 4.62. Meanwhile, 'The learning design shows that students are offered regular, structured and authentic feedback to improve the quality of their work' had the lowest score.

4. Discussion

This study aimed to investigate if ChatGPT could generate inquiry-based lesson plans for middle-school science courses in a Turkish context. To achieve this, an exploratory research design was adopted. The topic of 'Matter and its Nature' was chosen as the subject of the case. ChatGPT was asked to create lesson plans based on the objectives and duration of the lessons. Then, the plans were evaluated by the researchers considering the rubric developed by Borda-Carulla and Harmen (2012). According to the findings, there were both strengths and areas for improvement in AI-generated lesson plans.

The AI-generated lesson plans showed a high degree of alignment with learning outcomes ($M=4.81$) and the promotion of inquiry ($M=4.29$). This may indicate that the activities that were generated by ChatGPT were effectively designed to achieve the objectives and were accessible to users. Furthermore, it reflects that the lesson plans were successful in encouraging students to investigate and explore. This means that AI can generate educational materials and design lessons considering educational outcomes for an inquiry-based learning environment. This finding is consistent with the relevant literature. For example, studies by Cooper (2023) have shown that AI systems can design lesson plans that adhere to curriculum goals and promote inquiry, albeit with variations depending on the AI's training and context of use.

The moderate score in real-world connections ($M=4.14$) and student engagement ($M=3.00$) highlights a gap, which is also reflected in some national and international research. In Türkiye, research such as Çelik (2019) suggests that while AI-generated content can be contextually relevant, it often lacks the depth needed to deeply engage students with real-world applications. Similarly, research by Hattie (2009) emphasizes that engaging students in real-world contexts is crucial for effective learning, and AI systems may need more sophisticated mechanisms to enhance this aspect. Furthermore, the moderate success in promoting collaborative learning ($M=4.1$) is consistent with findings in the literature that AI can facilitate collaborative environments, but there is room for improvement. This suggests that AI-generated lesson plans may require enhanced features to foster more meaningful and productive collaborative experiences. Integrating more advanced collaborative tools and real-world scenarios into AI-generated content could help bridge these gaps and better support rich learning experiences.

Another dimension in which the plans were successful was documentation ($M=4.38$). In other words, the lesson plans generally emphasized that students recorded and presented their findings, which is also an essential skill of scientific inquiry. The dimension of evaluation assessed the evaluation of the activity. The mean score of this dimension was 3.62, which suggests that the lesson plans provided moderate opportunities to evaluate the activities and needed improvement, perhaps by adding reflective activities. Meanwhile, the lowest score for the general dimensions concerned feedback ($M=2.86$) and indicated that there was a need for improvement in the lesson plans regarding providing feedback to students. The low score suggests the importance of improving AI to integrate more comprehensive feedback mechanisms into plans while designing an inquiry-based science course.

To sum up, the lesson plans generated by ChatGPT generally received moderate or high scores from the Guidelines for Assessing Inquiry-based Lesson Plans rubric. Aligning with learning outcomes, encouraging real-world connections, promoting inquiry, and encouraging collaboration and

documentation were the strengths of the lesson plans. These features suggest that such lesson plans could aid teachers in creating effective inquiry-based learning environments. It is also important to consider that there are studies which suggest a positive relationship between inquiry-based learning and students' positive outcomes (e.g., Suduc et al., 2015; Ünlü, 2015). However, integrating inquiry-based learning effectively is crucial, and teachers need support in this area (Dobber et al., 2017; Sarıoğlu et al., 2016; Cleovoulou & Beach, 2019). The findings of this study support that AI has the potential to develop inquiry-based lesson plans. Nevertheless, the provision of feedback and the inclusion of more starting points for curiosity are areas for improvement. Integrating feedback into students' work was the weakest part of the lesson plans. It is important to note that ChatGPT was not specifically trained in the current study. In future studies, it may be helpful to train an AI model to generate more rounded plans.

References

- Açıkgöz, D., Sağır, Ş. U., & Ozan, F. (2018). Araştırma sorgulamaya dayalı öğretim hakkında öğretmen tutumları ölçek uyarlama çalışması. *Uluslararası Türk Eğitim Bilimleri Dergisi*, 6(11), 11-14.
- Barrow, L. H. (2006). A brief history of inquiry: From Dewey to standards. *Journal of Science Teacher Education*, 17(3), 265-278. <https://doi.org/10.1007/s10972-006-9008-5>
- Borda-Carulla, S., & Harmen, W. (2012). Tools for enhancing inquiry in science education. *Fibonacci Project*.
- Brennan, R. L., & Prediger, D. J. (1981). Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41(3), 687-699. <https://doi.org/10.1177/001316448104100307>
- Can, G. (1998). Fen bilgisi öğretiminde ölçme ve değerlendirme. Ş. Yaşar (Ed.), *Fen bilgisi öğretimi*.
- Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Cleovoulou, Y., & Beach, P. (2019). Teaching critical literacy in inquiry-based classrooms: Teachers' understanding of practice and pedagogy in elementary schools. *Teaching and Teacher Education*, 83, 188-198. <https://doi.org/10.1016/j.tate.2019.04.012>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.).
- 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>
- Cooper, J. M. (2011). The effective teacher. In J. M. Cooper (Ed.), *Classroom teaching skills* (9th ed.).
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.
- Cullern, K. A. (2014). *Leading curriculum change: Developing inquiry-based teaching and learning in a primary school* [Doctoral dissertation, University of Nottingham].
- Çelik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, Article 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Çelik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66(4), 616-630. <https://doi.org/10.1007/s11528-022-00715-y>
- Degenhart, H. S. (2010). Relationship of inquiry-based learning elements on changes in middle school students' science, technology, engineering, and mathematics (STEM) beliefs and interests [Doctoral dissertation, Texas A & M University].
- Dobber, M., Zwart, R., Tanis, M., & van Oers, B. (2017). Literature review: The role of the teacher in inquiry-based education. *Educational Research Review*, 22, 194-214. <https://doi.org/10.1016/j.edurev.2017.09.002>
- Fibonacci. (2012). *Inquiry in mathematics education*. <http://fibonacci.uni-bayreuth.de/resources/resources-for-implementing-inquiry.html>
- Güler, B., & Sahin, M. (2018). The inquiry-based learning from pre-service science teachers' perspective. *Kastamonu Eğitim Dergisi*, 26(5), 1561-1569. <https://doi.org/10.24106/kefdergi.2146>

- Hamilton, M. L., Smith, L., & Worthington, K. (2008). Fitting the methodology with the research: An exploration of narrative, self-study and auto-ethnography. *Studying Teacher Education*, 4(1), 17–28. <https://doi.org/10.1080/17425960801976321>
- Havuz, C., & Karamustafaoğlu, S. (2016). Fen bilgisi öğretmen adaylarının araştırma-sorgulamaya dayalı öğrenme algılarının incelenmesi. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 5(1), 233–247.
- Hüner, S. B. (2018). *Sokratik sorgulama temelli etkinliklerin hayat bilgisi dersinde başarı ve kalıcılığa etkisinin incelenmesi: Bir eylem araştırması* [Doktora tezi, İstanbul Üniversitesi].
- Inoue, N., Asada, T., Maeda, N., & Nakamura, S. (2019). Deconstructing teacher expertise for inquiry-based teaching: Looking into consensus building pedagogy in Japanese classrooms. *Teaching and Teacher Education*, 77, 366–377. <https://doi.org/10.1016/j.tate.2018.10.016>
- Jerrim, J., Oliver, M., & Sims, S. (2019). The relationship between inquiry-based teaching and students' achievement. New evidence from a longitudinal PISA study in England. *Learning and Instruction*, 61, 35–44. <https://doi.org/10.1016/j.learninstruc.2018.12.004>
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2017). *The 2017 horizon report*. EDUCAUSE Learning Initiative.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2, 130–144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Kayacan, K., & Selvi, M. (2017). Öz düzenleme faaliyetleri ile zenginleştirilmiş araştırma-sorgulamaya dayalı öğretim stratejisinin kavramsal anlamaya ve akademik öz yeterliğe etkisi. *Kastamonu Eğitim Dergisi*, 25(5), 1771–1786.
- Keçeci, G. (2014). *Araştırma ve sorgulama dayalı fen öğretiminin öğrencilerin bilimsel süreç becerilerine ve tutumlarına etkisi* [Doktora tezi, Fırat Üniversitesi].
- Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching*, 40(9), 898–921. <https://doi.org/10.1002/tea.10115>
- Khan, I. A. (2011). Lesson planning for teaching English. *Elixir Social Studies*, 37, 3958–3964.
- Krajcik, J. S., & Czerniak, C. M. (2014). *Teaching science in elementary and middle school: A project-based approach*. Routledge.
- Liang, Q., de la Torre, J., & Law, N. (2021). Do background characteristics matter in children's mastery of digital literacy? A cognitive diagnosis model analysis. *Computers in Human Behavior*, 122, Article 106850. <https://doi.org/10.1016/j.chb.2021.106850>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
- Millî Eğitim Bakanlığı. (2018). *Fen bilimleri dersi öğretim programı (İlkokul ve ortaokul 3, 4, 5, 6, 7 ve 8. sınıflar)*.
- Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment Research & Evaluation*, 7(10). <http://PAREonline.net/getvn.asp?v=7&n=10>
- National Research Council. (2000). *Inquiry and the national science education standards: A guide for teaching and learning*. The National Academies Press. <https://doi.org/10.17226/9596>
- OpenAI. (2023). *ChatGPT: Optimizing language models for dialogue*. <https://openai.com/blog/chatgpt/>

- Pender, H. L., Bohl, L., Schönberger, M., & Knopf, J. (2022). An AI-based lesson planning software to support competency-based learning. *International Conference on Higher Education Advances, 2022-June*, 1033–1041. <https://doi.org/10.4995/HEAd22.2022.14599>
- Saka, T. (2018). 5. Sınıf fizik konularının öğretimine yönelik rehberli araştırma sorgulamaya dayalı öğretmen kılavuz materyali geliştirilmesi ve değerlendirilmesi [Doktora tezi, Trabzon Üniversitesi].
- Sarıoğlu, A. B., Can, Y., & Gedik, İ. (2016). 6. Sınıf fen bilimleri ders kitabındaki etkinliklerin araştırma-sorgulamaya dayalı öğrenme yaklaşımına uygunluğunun değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 16(3), 1004–1025.
- Sergis, S., Sampson, D. G., Rodríguez-Triana, M. J., Gillet, D., Pelliccione, L., & de Jong, T. (2019). Using educational data from teaching and learning to inform teachers' reflective educational design in inquiry-based STEM education. *Computers in Human Behavior*, 92, 724-738. <https://doi.org/10.1016/j.chb.2017.12.014>
- Suduc, A. M., Bizoi, M., & Gorghiu, G. (2015). Inquiry-based science learning in primary education. *Procedia-Social and Behavioral Sciences*, 205, 474-479. <https://doi.org/10.1016/j.sbspro.2015.09.044>
- Şahingöz, S. (2017). *An investigation of Turkish middle school science teachers' pedagogical orientations towards direct and inquiry instructional approaches* [Doctoral dissertation, Western Michigan University].
- Teig, N., Scherer, R., & Nilsen, T. (2018). More isn't always better: The curvilinear relationship between inquiry-based teaching and student achievement in science. *Learning and Instruction*, 56, 20–29. <https://doi.org/10.1016/j.learninstruc.2018.02.006>
- Thaiposri, P., & Wannapiroon, P. (2015). Enhancing students' critical thinking skills through teaching and learning by inquiry-based learning activities using social network and cloud computing. *Procedia-Social and Behavioral Sciences*, 174, 2137-2144.
- UNESCO. (2024). *AI competency framework for students*. United Nations Educational, Scientific and Cultural Organization. <https://doi.org/10.54675/JKJB9835>
- Ünal, A. (2018). *Araştırma-sorgulamaya dayalı ve sosyal ağ destekli kimya laboratuvarı etkinliklerinin fen bilimleri öğretmen adaylarının algı, tutum ve başarıları üzerine etkisi* [Doktora tezi, Kastamonu Üniversitesi].
- Ünal, Ş. (2015). *Biyoloji dersi çevre konularının öğretiminde yaşam temelli yaklaşıma dair örnek olay inceleme ve araştırma sorgulama temelli bilim öğrenme öğretim yöntemlerinin etkisi* [Doktora tezi, Atatürk Üniversitesi].
- Van den Berg, G., & du Plessis, E. (2023). ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*, 13(10), Article 998. <https://doi.org/10.3390/educsci13100998>
- Varlu, B., & Sağır, Ş. U. (2019). Araştırma sorgulamaya dayalı öğretimin ortaokul öğrencilerinin fen başarıları, sorgulama algısı ve üstbiliş farkındalığına etkisi. *GÜJGEF*, 39(2), 703–725. <https://doi.org/10.17152/gefad.407417>
- Villardón-Gallego, L. (2016). Inquiry-based learning in pre-service training for secondary education counselors. *Procedia-Social and Behavioral Sciences*, 217, 65–73. <https://doi.org/10.1016/j.sbspro.2016.02.028>

- Yaşar, Ş., & Duban, N. (2009). Students' opinions regarding to the inquiry-based learning approach. *Elementary Education Online*, 8(2), 457–475. <https://ilkogretim-online.org/index.php/pub/article/view/2469>
- Yerlikaya, A. (2019). *7. sınıf vücudumuzdaki sistemler ünitesinin araştırma sorgulama temelli öğreniminin programdaki öğrenme alanlarına etkisi* [Doktora tezi, On Dokuz Mayıs Üniversitesi].
- Yetişir, M. I. (2016). Rehberli araştırma-sorgulamaya dayalı fizik öğretimi: Öğretmen adaylarının akademik başarıları ve uygulama hakkındaki görüşlerinin incelenmesi. *Ankara University Journal of Faculty of Educational Sciences*, 49(1), 159–182. https://doi.org/10.1501/Egifak_0000001379
- Yıldırım, E., & Yıldırım, Ö. (2020). Investigation of lesson planning competencies of primary and secondary school teachers. *Milli Eğitim*, 49(228), 7–37.
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1). <https://doi.org/10.1186/s41239-019-0171-0>

Article Information Form

Authors Contributions: Nurcan Kahraman design of the study, data analysis, and writing, Gülbin Kıyıcı data analysis, and writing.

Conflict of Interest Disclosure: No potential conflict of interest was declared by authors.

Artificial Intelligence Statement: No artificial intelligence tools were used while writing this article. This study investigated AI-generated lesson plans, and ChatGPT was used in the generation of AI-based lesson plans. However, no AI tools were used while writing the manuscript.

Plagiarism Statement: This article has been scanned by iThenticate.