

## Research Article

# A Non-AI Homework Support Tool to Enhance Achievement and Interest in Science Education: BilgeCan Bot

Taner Yılmaz<sup>1a</sup>, İlbilge Dökme<sup>2b</sup><sup>1</sup>Medical Services and Techniques Department, Cekerek Fuat Oktay Vocational School of Health Services, Yozgat Bozok University, Yozgat, Türkiye<sup>2</sup>Science Education Department, Education Faculty, Gazi University, Ankara, Turkey

taner.yilmaz@bozok.edu.tr

DOI : 10.31202/ecjse.1726783

Received: 25.06.2025 Accepted: 08.07.2025

**How to cite this article:**

Taner Yılmaz, İlbilge Dökme, " A Non-AI Homework Support Tool to Enhance Achievement and Interest in Science Education:

BilgeCan Bot ", El-Cezeri Journal of Science and Engineering, Vol: 12, Iss: 3, (2025), pp.(249-273).

ORCID: <sup>a</sup> 0000-0003-1164-3549, <sup>b</sup>0000-0003-0227-6193.

**Abstract** Homework helps students learn and develop independent study skills, but they often need extra guidance and reliable information sources, especially in science where abstract concepts can be challenging. Traditional resources may not match students' cognitive levels and can lead to information overload. To address this, non-AI rule-based (NARB) educational chatbots can provide focused, essential information with minimal cognitive load. This study explores the effects of BilgeCan Bot, a NARB chatbot designed to help middle school students with astronomy homework, based on Cognitive Load Theory. Conducted during the 2022–2023 academic year with 52 fifth-grade students in Türkiye, the research used an explanatory sequential mixed-methods approach. Students were divided into an experimental group using BilgeCan Bot and a control group using textbooks. Data was collected through tests, interest scales, interviews, chatbot records, and teacher feedback. Results showed that students using BilgeCan Bot achieved higher scores and had greater interest in science. They described the chatbot as effective and engaging, helping them grasp difficult concepts. Overall, the findings suggest that NARB chatbots can offer targeted, reliable support for homework in science education without overwhelming students.

**Keywords:** Non-AI Rule-Based educational chatbot, Science education, Academic achievement, Cognitive load theory, Student interest.

## I. INTRODUCTION

Chatbots are software programs that simulate human-like conversations through pattern recognition techniques and are increasingly recognized for their educational potential [1]. Educational chatbots (ECs) employ natural language processing algorithms to provide personalized guidance, facilitate both in-class and extracurricular learning, optimize students' learning pace, and enhance their understanding of abstract concepts [2–4]. Chatbots are typically categorized as rule-based or AI-based systems [5]. While AI-based chatbots adapt to student input, rule-based chatbots deliver information through predefined scenarios, offering structured guidance that mitigates cognitive load [5–8]. Particularly in science education, where abstract concepts prevail, structured information delivery enhances learning efficiency and academic achievement [9, 10].

Homework is a critical tool for assessing student progress, reinforcing classroom learning, and fostering independent study skills [11, 12]. When structured support is lacking at home, students seek reliable resources, particularly when facing conceptual difficulties [13–15]. This need is pronounced in science education, which demands abstract thinking and often leads to misconceptions, especially within topics like the "Sun, Earth, and Moon" unit [16, 17]. If unaddressed, these misconceptions may persist across educational stages [18]. Therefore, accurate instruction in abstract scientific concepts is essential for cognitive development [19]. Additionally, challenges such as time management and motivation during homework further highlight the necessity of structured support systems [20].

Students frequently use search engines such as Google and Yahoo to address homework-related challenges; however, these tools often provide unfiltered, context-independent information that limits their pedagogical utility [21, 22]. Consequently, the demand for dialogue systems offering meaningful, natural language feedback has grown [23, 3]. Educational chatbots meet this need by delivering real-time responses, fostering interactive learning, and enhancing students' motivation and depth of understanding [24]. Studies further indicate that educational chatbots strengthen teacher-student interaction and improve the effectiveness of homework support [25, 26]. The use of generative AI (GAI) systems such as ChatGPT for information retrieval among students has increased; however, their pedagogical reliability remains contested. Research indicates that students often struggle to assess online content credibility, with only 7% correctly identifying a fake website [27]. Given that GAI systems frequently draw from anonymous sources, they pose risks of misinformation and cognitive overload, potentially impairing critical thinking and offering contextually inappropriate content [28, 29].

In this context, rule-based educational chatbots facilitate learning by delivering only essential information, thereby reducing excessive cognitive load [30]. According to Cognitive Load Theory (CLT), effective learning requires optimizing intrinsic, extraneous, and germane cognitive loads [31, 32]. While AI-based chatbots may increase extraneous or germane loads through unstructured information, rule-based chatbots mitigate these risks by offering structured, focused content. Consequently, empirical research is needed to assess their effectiveness in teaching abstract subjects such as astronomy. BilgeCan Bot, developed in this study, operates on predefined rules without incorporating machine learning or deep learning techniques, and is thus classified as a non-AI rule-based (NARB) chatbot [33, 23, 34]. Although the effects of AI-supported chatbots on academic achievement and science interest have been widely studied [35, 36], limited research addresses the role of NARB chatbots in guiding homework tasks [37]. Particularly in middle school science education, the impact of NARB chatbots on student achievement and interest remains underexplored. This study aims to evaluate the effectiveness of BilgeCan Bot in supporting middle school students' astronomy homework by providing only necessary information and minimizing cognitive load.

## II. BACKGROUND AND RELATED WORK

### 2.1. Chatbot Design Approaches and Rule-Based Chatbots

The term "chatbot" was first introduced by Mauldin in 1997 [38] and has since been known by terms like "Conversational Agent" and "Virtual Assistant" [39, 40]. Despite the variety, "chatbot" remains the most widely accepted term [41]. Early developments include the Turing Test and ELIZA, which used keyword-based transformation rules [33, 42].

Chatbots are classified based on purpose, domain, platform, and design [43–45]. Madhu et al. (2017) categorized them as AI-based (intelligent) and rule-based (non-intelligent). Rule-based systems, exemplified by ELIZA, are easier to implement and provide more reliable, structured responses [47], while AI-based systems may produce context-free or inaccurate outputs [16].

In education, chatbots like AutoTutor and Duolingo offer structured support [5]. Two main design approaches exist: pattern matching and machine learning [45]. Rule-based chatbots generate responses by matching inputs to predefined templates [48].

BilgeCan Bot, developed for this study, uses regex-based pattern matching without NLP or AIML [49], responding strictly based on predefined structures [50]. While "keyword matching" and "pattern matching" are often used interchangeably, deeper analysis is sometimes distinguished as "keyword recognition" [44, 43].

Most research focuses on AI-supported systems [51, 52], leaving a gap in understanding non-AI rule-based (NARB) chatbots for homework support. This study addresses that gap by exploring the educational potential of simplified NARB systems in middle school science education.

### 2.2. Cognitive Load Theory and Rule-Based Chatbots

Cognitive Load Theory (CLT) aims to optimize instructional design by considering the limited cognitive capacity of humans during the learning process [31]. According to the theory, three types of cognitive load occur during learning: intrinsic, extraneous, and germane load. CLT particularly emphasizes reducing extraneous load to allow more efficient use of students' cognitive resources [7]. In this context, it is important to design digital learning environments in a simple and guided manner [53]. Rule-based chatbots can help reduce extraneous cognitive load by providing only the necessary information and supporting guided learning [6]. Especially in science courses dealing with abstract concepts, they facilitate load management by delivering structured and directed content [54]. Compared to traditional sources, they offer more focused responses, minimizing unnecessary mental processing [55]. These systems can enhance learning efficiency by providing real-time feedback [56]. Since they operate based on predefined patterns without using natural language processing or artificial intelligence, they deliver information quickly, reliably, and with a low risk of error. When appropriately designed, they can optimize cognitive load by offering content tailored to individual needs [30] and promote active learning through interactive engagement [56].

In conclusion, rule-based chatbots developed within the framework of CLT have the potential to improve the learning experience by preventing information overload, providing guidance, and supporting personalized learning [57]. However, studies directly addressing the effects of rule-based chatbots on cognitive load management in science courses remain limited [6, 54]. This highlights the need for systematic investigation of systems like BilgeCan Bot in the context of science education.

### 2.3. The Use of Chatbots in Science Education

Science education often presents challenges for students due to the abstract nature of concepts and associated cognitive load [58–60]. Consequently, digital tools are increasingly used to support conceptual understanding and facilitate learning [61]. Among these tools, chatbots have emerged as effective by providing real-time guidance and step-by-step learning support.

Astronomy topics, being unobservable and highly abstract, are particularly difficult for students [62, 63]. Studies show that students often struggle with basic concepts such as the solar system, phases of the Moon, celestial movements, and seasonal changes [64, 65]. Interactive digital tools can help reduce misconceptions by making abstract ideas more concrete. Augmented reality, for example, visualizes celestial movements [5], while chatbots offer real-time responses to student inquiries, guiding conceptual understanding [66].

The interdisciplinary nature of astronomy demands flexible learning tools, as rigid curricula may restrict inquiry-based learning [63]. Chatbots can bridge this gap by explaining scientific concepts, providing visual support, and offering personalized learning opportunities [61]. Rule-based systems like BilgeCan Bot enhance learning efficiency by delivering structured guidance [67].

Research within the framework of Cognitive Load Theory (CLT) shows that material presentation directly affects conceptual understanding [57], and excessive cognitive load hampers learning [68]. Digital guides and chatbots can balance cognitive load by delivering targeted, essential information [32]. Instant feedback and structured guidance help internalize scientific concepts. Overall, chatbots hold strong potential for improving conceptual understanding, academic achievement, and scientific interest, especially in abstract fields like astronomy [69].

#### 2.4. Homework, Academic Achievement, Interest in the Subject, and Chatbot-Supported Learning

Homework reinforces classroom learning, fosters independent study, and boosts academic achievement [70]. However, in abstract subjects like science, students often need additional structured guidance [71]. Stimulating curiosity and engagement is critical for effective science learning [72], and digital tools offering structured support play a key role [73].

Research shows that regular, structured homework enhances retention, self-regulation, and long-term learning [74, 75]. Yet challenges like conceptual misunderstandings, unreliable information access, and lack of timely guidance persist [76]. Traditional resources may lead to information overload [55].

Interactive digital tools providing real-time, structured support ease learning and reduce cognitive load [77, 78]. In this context, chatbots offer targeted information efficiently, improving both learning and motivation [54, 79]. They foster active participation and sustain engagement [36, 16, 3].

Rule-based chatbots, in particular, deepen understanding through structured guidance [80–82]. Research identifies three key benefits of chatbot use in homework: correcting misconceptions [61], providing immediate feedback [116], and enhancing self-regulation [82]. Overall, chatbot-supported environments improve learning efficiency, reduce cognitive overload, and boost achievement and motivation, especially in science education [6, 83, 84].

### III. AIM OF THE STUDY

Non-AI rule-based (NARB) chatbots can prevent information overload by providing only targeted information, allowing students to experience a focused and guided learning process [67]. In this context, a non-AI rule-based chatbot named BilgeCan Bot was developed to support middle school students with their astronomy-related homework. The aim of this study is to examine the effects of BilgeCan Bot on students' academic achievement in science courses and their interest in science. Additionally, by analyzing students' perceptions of chatbots, the study seeks to offer insights into how such guidance systems can be used more effectively in educational settings. Based on these aims, the following research questions are addressed:

- RQ1: When controlling for pre-test scores, is there a significant difference in post-test academic achievement and interest in science between students who used BilgeCan Bot (experimental group) and those who used the textbook (control group) while completing homework?
- RQ2: What are the perceptions of the experimental group students regarding BilgeCan Bot after the experimental implementation?
- RQ3: What is the nature of the conversations conducted by the experimental group students with BilgeCan Bot?
- RQ4: What challenges did the control group students face while completing their homework using the textbook?

### IV. DESIGNING A HOMEWORK SUPPORT CHATBOT SYSTEM (BILGECAN BOT)

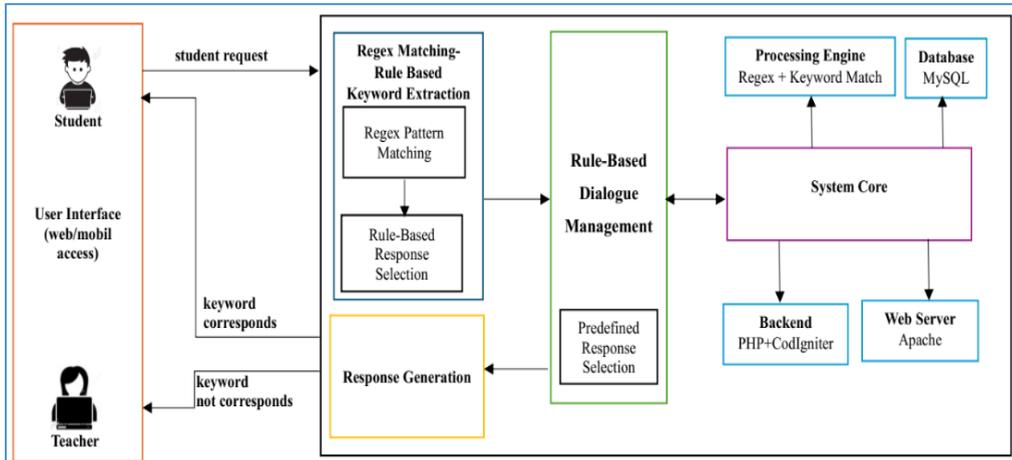
#### 4.1. Overview of BilgeCan Bot

BilgeCan Bot is a web-based, rule-based chatbot system developed to provide homework support for middle school science courses. Designed based on Cognitive Load Theory, the system aims to prevent excessive cognitive load and facilitate learning by offering students only the information they need. Developed with a JavaScript-based front-end, the system can be accessed via an internet browser on either a computer or a mobile device (Figure 1). Users can register by selecting the “Create New Account” option and then log in using their email and password to access the system.

Figure 1. BilgeCan bot login

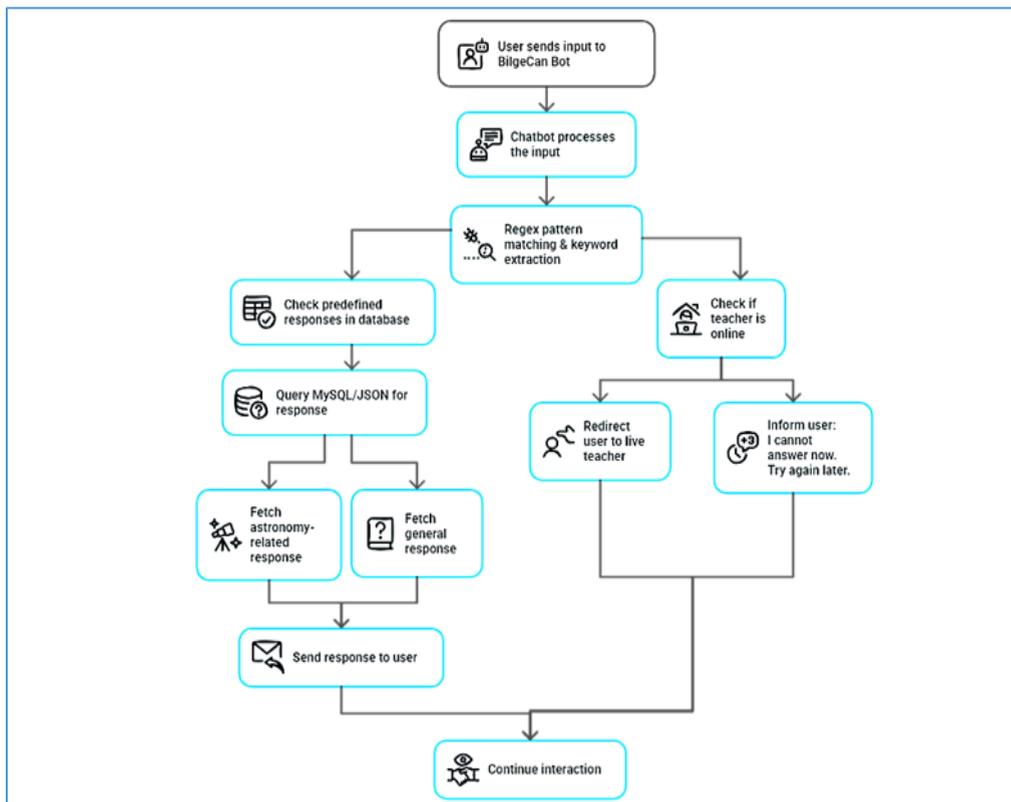
**4.2. The Architecture and Operational Process of BilgeCan Bot**

BilgeCan Bot’s architecture consists of three components: data preparation, user interaction, and technological infrastructure (Figure 2). Learning materials are structured into JSON/XML formats via regex-based keyword extraction. Developed with PHP 7.4 and CodeIgniter 3, the system stores data in a MySQL 10.6 database. Students interact with the bot through a text-based web interface, with requests processed by an Apache 2.4 server. Using regex and keyword matching, BilgeCan Bot operates as a lightweight, fast, and efficient rule-based learning assistant.



**Figure 2. The system architecture of BilgeCan bot**

BilgeCan Bot analyzes student questions using regex rules, identifies keywords, and retrieves appropriate astronomy-related answers from its database. If no answer is found, the bot checks teacher availability: if online, the student is directed to the teacher; if not, a notification is displayed. The bot ensures continuity by responding to subsequent questions (Figure 3).



**Figure 3. Process of response generation and operation in bilgecan bot**

#### 4.3. Dialogue Mechanism of BilgeCan Bot

BilgeCan Bot's dialogue system utilizes RegEx (Regular Expressions). RegEx is a powerful tool for recognizing and matching patterns in strings and is widely used in text processing applications. A similar approach was also employed in the first chatbot, ELIZA [85]. RegEx defines patterns by combining normal and special characters (\*, +, ?), allowing the parsing and processing of user inputs [86]. Chatbots generate responses by matching user statements with predefined patterns. Wildcard characters make the patterns more flexible, enabling appropriate responses to questions phrased in different ways [87]. BilgeCan Bot is capable of providing flexible and effective responses using this method. An example question constructed with wildcard characters is shown in Figure 4.

**Soru**

T(ü|u)?rki(y|q)e?(n)?i(n)?( '|')?ba(s|ş)(k|q)en(t)?(i)?( '|')?n(e|w)re(si)?

Figure 4. A Sample of wildcards used in the bilgecan bot

The example shown in Fig. 4 is designed to recognize different variations and errors in phrasing the question "What is the capital of Türkiye?" Thanks to the use of wildcard characters, variations such as "What is the capital of Türkiye" or "What is the capital of Türkiye's" can still be matched with the correct response (Figure 5).

```
import re
# Regex ile eşleşen soru desenleri
regex_pattern = r"T(ü|u)?rki(y|q)e?(n)?i(n)?( '|')?ba(s|ş)(k|q)en(t)?(i)?( '|')?n(e|w)re(si)?"
# Kullanıcının sorduğu soruların farklı varyasyonları
test_questions = [
    "Türkiye'nin başkenti neresi?", # Doğru yazım
    "Türkiye'nin başkenti neresi?", # Ü harfi yerine u kullanılmış
    "Türkiyenin başkenti neresi?", # Büyük/küçük harf farkı
    "türkiya'nın başkenti neresi?", # Yazım hataları içeriyor
    "Türkiyenin başkenti ner'si?", # Eksik harfler ve farklı yazım
    "Tuerkiyenin başkenti neresi?" # Q harfi kullanılmış
]
# Kullanıcının girdilerini regex ile kontrol etme
for question in test_questions:
    if re.search(regex_pattern, question, re.IGNORECASE):
        print(f"Eşleşme bulundu: {question}")
    else:
        print(f"Eşleşme bulunamadı: {question}")
```

Figure 5. Regex-Based keyword matching code structure

#### 4.4. User Interface Features of BilgeCan Bot

The teacher interface of BilgeCan Bot consists of two sections: system management and content control. The first section, shared with the student interface, includes menus such as "User Settings," "Chat," "Messages," "Active Teachers," and "User Guide," allowing teachers to monitor student activity (detailed in the "Student Interface Features" section). The second section, focusing on content and system management, includes the "Questions," "Surveys," "Tests," "System," "Feedback," and "Active Students" menus for content addition, user tracking, and system operations (Figure 6).

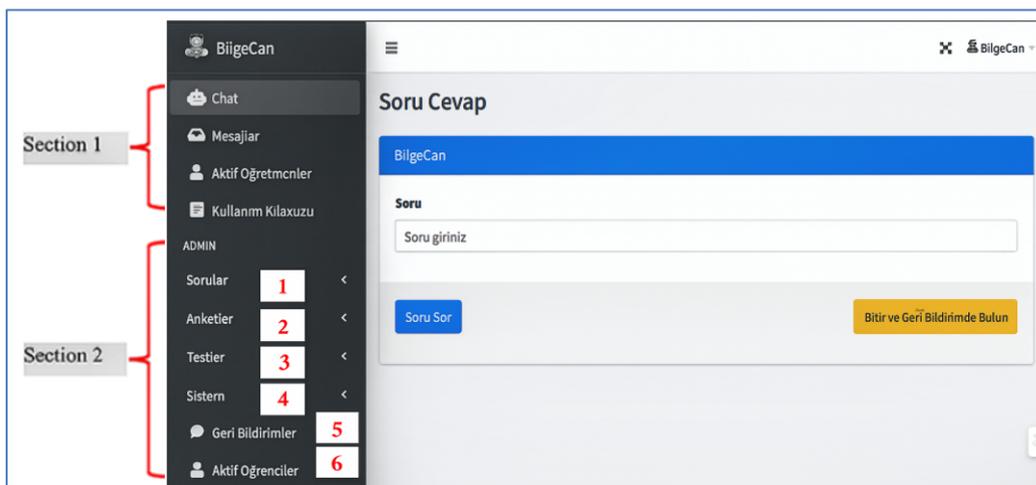


Figure 6. Teacher interface

The second section of the teacher interface includes tools for managing BilgeCan Bot's content and operations. The Questions menu (1) allows teachers to manage inquiries and add new content manually or via CSV files. The Surveys menu (2) enables survey creation and result review, planned for future use. The Tests menu (3) supports the creation of open-ended or multiple-choice assessments. The System menu (4) manages user registrations, roles, and user guide updates. The Feedback menu (5) collects student opinions for system improvement. The Active Students menu (6) allows real-time monitoring of online student activity.

The student interface includes menus for settings, notifications, chat, messages, test results, and active teachers, all supporting the learning process (Fig. 7). In the Chat menu (1), students can ask questions to BilgeCan Bot; if the bot cannot answer, they are directed to an online teacher. The Messages menu (2) displays past conversations. The Active Teachers menu (3) shows which teachers are online. The User Guide menu (4) provides usage instructions. The Finish and Provide Feedback menu (5) allows students to end the session and give feedback. The Enter Your Question field (6) is used to input messages.

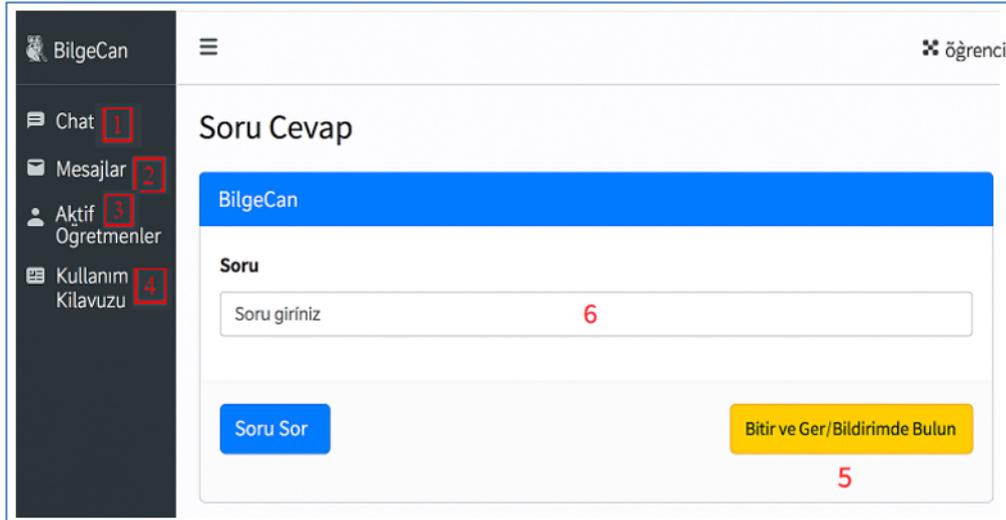


Figure 7. Student interface

#### 4.5. Pilot Implementation with Bilgecan Bot and Preparation of Dialogues

Before the main study, a pilot implementation was conducted with sixth-grade students to test the usability of BilgeCan Bot and enhance student-teacher interaction. The bot was introduced, user registrations were completed, and interactions were maintained for one week. At the end of the process, it was determined that the bot supported conceptual understanding, its language was simplified, and some explanatory texts were added. In addition, short supportive phrases such as greetings, thanks, and guidance were incorporated to establish a friendly and guiding communication with students (Figure 8).



Figure 8. Samples of auxiliary text

The main texts are focused on three key topics from the “Sun, Earth, and Moon” unit of the fifth-grade science curriculum: “The Structure and Features of the Sun,” “The Structure and Features of the Moon,” and “The Movements and Phases of the Moon.” These texts were prepared in accordance with the learning outcomes specified in the 2018 TMoNE (Turkish Ministry of National Education) curriculum. For example, dialogues were developed to address objectives such as explaining the features of the Sun and comparing the size of the Sun with that of the Earth (Figure 9).

<p><b>Soru</b></p> <p>Neden ateş topuna benzediğini bilmiyorum</p> <p><b>Cevap</b></p> <ul style="list-style-type: none"> <li>Güneş'in rengini Güneş ışınlarının gözümüze geliş açısından dolayı farklı algılarız.</li> <li>Güneş, orta büyüklükte bir yıldızdır ve yıldızların rengi sıcaklıklarına bağlı olarak değişir.</li> <li>Güneş'in içinde bulunan Helyum Hidrojene dönüşerek etrafına ısı ve ışık yayar.</li> </ul> <p>Bu yüzden Güneş'in rengi ateş topuna benzer.</p>  <p>❓ Peki sence Güneş mi daha büyük yoksa Dünya mı?</p> <p>Eğer bu sorunun cevabını bilmiyorsan soru sorma kutucuğuna "Bu soru hakkında hiçbir fikrim yok" şeklinde yaz.</p> <p>Eğer bu sorunun cevabını biliyorsan soru sorma kutucuğuna "Bu sorunun cevabını biliyorum" şeklinde yaz.</p>	<p><b>Soru</b></p> <p>Neden ateş topuna benzediğini biliyorum</p> <p><b>Cevap</b></p> <p>🤖 Ama yine de cevabı görmek istersen " Neden ateş topuna benzediğini bilmiyorum" yaz. Senin bilginle benim bilgilerimi karşılaştırmanı tavsiye ederim 🙌</p> <p>Belki de ikimizin cevapları birbirinden farklı olabilir 🤔</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 9. A sample of dialogue on the topic “the structure and properties of the sun”

Dialogues were created for the topics “The Structure and Features of the Moon” and “The Movements and Phases of the Moon” to support students’ understanding of the Moon’s characteristics, potential for life, and its movements (Figure 10). The dialogues encouraged students to assess their knowledge and compare it with the bot’s responses. Regardless of their confidence, students were prompted to review key concepts, with clear explanations provided when needed. Alternative phrasing allowed students to access information through differently worded questions. Guiding questions were also included to enhance engagement and promote natural learning.

<p><b>Soru</b></p> <p>Bu sorunun cevabını öğrenebilir miyim</p> <p><b>Cevap</b></p> <p>Ay, Dünya'nın doğal uydusudur</p> <p>Ay'a ilk ayak basan astronot Neil Armstrong'dur. 16 Temmuz 1969'da Apollo-11 uzay aracı ile çıkmış olduğu Ay yolculuğu 3 gün sürmüştür.</p> <p>Gece ile gündüz arasındaki sıcaklık farkı çok fazladır. Gündüz sıcaklığı 107 °C'a kadar çıktığı, gece sıcaklığı ise -153 °C'a kadar düştüğü tespit edilmiştir. Bu yüzden Ay'da kayalar çok fazla parçalanarak toz ve kum oluşmasına neden olur. Ay'ın atmosferi yok denemek kadar az olduğu için meteor adı verilen gök cisimleri Ay'ın yüzeyine düşerek "krater" adı verilen çukurları oluştururlar.</p> <p>Güneş ve Dünya gibi Ay'ın şekli de küreye benzer. Yani Ay'ın şekli bir top gibidir.</p>  <p>👉 Geceleri gökyüzüne baktığımızda Ay'ın parlak olduğunu ve dünyamızı aydınlatıldığını görürüz.</p> <p>❓ Sence Ay, bir ışık kaynağı mı?</p> <p>Eğer bu sorunun cevabını bilmiyorsan "Bu soruya bir cevap veremeyeceğim" şeklinde yaz.</p> <p>Eğer bu sorunun cevabını biliyorsan "Bu soruya cevap verecek bilgin var" şeklinde yaz.</p>	<p><b>Soru</b></p> <p>Bu sorunun cevabını öğretmenimden öğrenmişim</p> <p><b>Cevap</b></p> <p>🤖 Ama yine de cevabı görmek istersen " Bu sorunun cevabını öğrenebilir miyim?" yaz. Senin bilginle benim bilgilerimi karşılaştırmanı tavsiye ederim 🙌</p> <p>Belki de ikimizin cevapları birbirinden farklı olabilir 🤔</p>
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 10. A sample of a dialogue on “the structure and properties of the moon and the movements and phases of the moon”

## V. METHOD

### 5.1. Research Design

An explanatory sequential (quantitative → qualitative) mixed methods design was used [88]. In the quantitative phase, a pre-test–post-test quasi-experimental design with a control group investigated RQ1 [89], focusing on differences between adjusted post-test means based on pre-test scores. In the qualitative phase, a phenomenological design explored RQ2, RQ3, and RQ4 [90]. Data were collected through semi-structured interviews (Appendix C), BilgeCan Bot conversation records, and control group teacher interviews. Sequential data collection strengthened the interaction between findings, with priority given to quantitative data for systematic analysis and qualitative data for in-depth interpretation.

### 5.2. Participants

A total of 52 fifth-grade students from a public middle school in Turkey participated in this study. The physical conditions of the classrooms were identical. Students were randomly assigned to experimental conditions. They were divided equally into two groups: the control group (Class A: 26 students; 10 boys and 16 girls) and the experimental group (Class B: 26 students; 12 boys and 14 girls). The students in both groups shared similar socioeconomic backgrounds.

### 5.3. Experimental Procedure

This study was conducted during the 2022–2023 academic year with fifth-grade students at a public middle school in Türkiye, within the "Sun, Earth, and Moon" unit. Necessary permissions were obtained from the university ethics committee, the Ministry of National Education, and parents through informed consent. Student confidentiality was ensured by assigning unique codes. The study followed the Turkish Ministry of National Education's 2018 science curriculum. The experimental process spanned seven weeks and involved BilgeCan Bot use, data collection, and analysis. Students in the experimental group used the bot outside of school to reinforce lesson topics, receiving predefined responses or being directed to teachers when needed. Teachers were available online on scheduled days, and class hour differences between groups were considered. Lessons for the experimental and control groups were conducted by different teachers using common lesson plans to minimize instructional differences. Classroom activities were monitored through checklists (Appendix A) and teacher interviews, though direct observation of interactions was not conducted. The experimental group teacher introduced BilgeCan Bot, provided usage instructions, and emphasized academic responsibilities. Students were encouraged to use the bot twice a week and to contact their teacher when necessary. Homework aligned with weekly learning objectives was assigned to both groups, with the experimental group supported by the bot and the control group using textbooks. The seven-week implementation schedule is summarized in Table 1.

The weekly implementation schedule was as follows:

- 1<sup>st</sup> Week: Pre-tests (AAT, SISL) administered; "Structure and Features of the Sun" covered.
- 2<sup>nd</sup> Week: Homework evaluation; "Comparison of Sun and Earth Sizes" and model-building activities.
- 3<sup>rd</sup> Week: "Structure and Features of the Moon" and scaling activities.
- 4<sup>th</sup> Week: "Movements and Phases of the Moon" with modeling, posters, and presentations.
- 5<sup>th</sup> Week: "Moon's Phases and Revolution" explored through interactive activities.
- 6<sup>th</sup> Week: "Sun, Earth, and Moon System" modeled; group work and presentations conducted.
- 7<sup>th</sup> Week: Post-tests (AAT, SISL) administered; semi-structured interviews conducted with the experimental group and teacher interviews conducted for the control group.

**Table 1. Procedure of the Implementation Period of the Research**

Week	In-school activities (teaching lessons)				Out-of-school activities (do homework)	
	Experimental and control groups				Experimental group	Control group
	Learning outcome	Learning activities	Teaching methods and techniques	Teaching materials	Use BilgeCan Bot	Use Textbook
	Pre-test and pre-questionnaire (achievement, interest)					
1st	Explains the properties of the Sun	"I will know the Sun better"	Expository teaching, question and answer technique, brainstorming, watching videos	Textbook, images, videos and animations	Worksheet-1	Worksheet-
2nd	Prepares a model to compare the Sun's and Earth's sizes	"We discover the size of the Earth and the Sun" "our earth and sun model"	Project based learning, brainstorming	Textbook, modeling materials	Worksheet-2	Worksheet-
3rd	Describes the size and geometric shape of the Moon	"I Will Know the Moon Better-1"	Expository teaching, question and answer technique, brainstorming, watching videos	Textbook, images, videos and animations	Worksheet-3	Worksheet-
4th	Explains the rotation and orbit movements of the Moon	"I Will Know the Moon Better-2"		Textbook, images, videos and animations	Worksheet-4	Worksheet-
5th	Explains the relationship between the phases of the Moon and the Moon's orbit around the Earth	"I Will Know the Moon Better-3"		Textbook, images, videos and animations	Worksheet-5	Worksheet-
6th	Prepares a model representing the motion of the Sun, Earth, and Moon relative to each other	"Our Earth and Moon Model"	Project based learning, Brainstorming Question and answer Technique, Watching videos	Modeling materials	Worksheet-6	Worksheet-
7th	Completion of posttests, semi-structured interviews and receive feedback					

#### 5.4. Measurements

Quantitative data in this study were collected using the Academic Achievement Test (AAT) and the Science Interest Scale (SIS). Qualitative data were gathered through a Semi-Structured Interview Form (SSIF) administered to the experimental group students after the intervention to collect their opinions about the chatbot. In addition, weekly conversation records between experimental group students and the chatbot, as well as feedback interviews with the control group's teacher regarding homework, were used as part of the qualitative data.

##### 5.4.1. Academic Achievement Test

An Academic Achievement Test (AAT) was developed for the "Sun, Earth, and Moon" unit of the fifth-grade science course within the scope of this study. The test covers the topics "The Structure and Features of the Sun," "The Structure and Features of the Moon," and "The Movements and Phases of the Moon." Items were selected from various sources based on the learning outcomes specified in the TMoNE (2018) Science Curriculum. Initially, 40 items were prepared, and content validity was ensured using a table of specifications. The test was reviewed by an academician, two science teachers, and one Turkish language teacher, and revised according to expert feedback. In the pilot implementation, three items were eliminated, and the remaining 37 items were analyzed using the TAP software. An additional 10 items were removed based on difficulty and discrimination indices, resulting in a final test with 27 items. Final analyses showed a mean item difficulty of 0.63, an item-total correlation of 0.43, and a KR-20 reliability coefficient of 0.88. These values indicate that the test is both valid and reliable [91–95].

##### 5.4.2. Scale of Interest in Science Lesson

In this study, the 27-item, 5-point Likert-type "Scale of Interest in Science Lesson (SISL)," developed by Laçın Şimşek and Nuhoglu (2009), was used to determine the science interest levels of fifth-grade students regarding educational chatbots. The scale consists of six sub-dimensions: exploring nature, understanding cause-and-effect relationships, conducting observations, relating concepts to daily life, following mass media, and individual interest. Participants can score between 27 and 135 points on the scale. The Cronbach's alpha reliability coefficient was reported as 0.79 in the original development study, while it was calculated as 0.74 for the pre-test and 0.78 for the post-test in this research. These results indicate that the scale is reliable [91]. Moreover, the findings are consistent with similar studies in the literature [97–99]. However, due to the limited sample size (N=52), a content validity analysis could not be performed, which is noted as a limitation of the study.

##### 5.4.3. Semi-Structured Interview Form

This form included 13 open-ended questions exploring students' experiences, expectations, and opinions about chatbot use (see Appendix B). Topics covered previous chatbot experiences, difficulties, liked and disliked features, improvement suggestions, and impacts on learning. The chatbot's effects on learning, questioning, thinking skills, attitudes toward science, and willingness to use it in other subjects were also evaluated. Its potential to support independent learning at home was examined. Student responses were analyzed using content analysis, with codes, categories, and themes developed. Reliability was ensured using Miles and Huberman's formula [100], with an inter-rater agreement of 86.9%, exceeding the 80% acceptability threshold.

##### 5.4.4. Weekly Chat Records of Experimental Group Students with BilgeCan Bot

The weekly conversation records between experimental group students and BilgeCan Bot were analyzed to identify students' interaction patterns with the chatbot, their information-seeking behaviors, and the challenges they encountered. In this process, both descriptive and thematic analysis methods were used to examine students' learning behaviors and the chatbot's contributions to the educational process. To assess the reliability of the analysis, inter-coder agreement was calculated, and consistency across the coding of student conversation records was determined. According to the classification scale for kappa values proposed by Landis and Koch [101], the inter-rater agreement was  $\kappa = 0.85$ , 95% CI [0.55, 1.12],  $p < .001$ , indicating an almost perfect level of agreement between coders. Sample conversations between students and both BilgeCan Bot and the live teacher are presented in Appendix C.

##### 5.4.5. Homework feedback interviews with the control group teacher

Regular interviews were conducted with the teacher responsible for the control group to discuss the difficulties students faced while completing their weekly homework assignments. Feedback collected from students by the teacher was analyzed using thematic analysis. To assess the reliability of the analysis, inter-coder agreement among three coders was calculated, and the consistency of thematic coding processes was determined. Krippendorff's alpha coefficient was calculated as  $\alpha = 0.48$ , indicating a moderate level of agreement among raters [102].

## 5.5. Data analysis

### 5.5.1. Analysis of Quantitative Data

In this study, both quantitative and qualitative data analysis methods were used to examine the effects of BilgeCan Bot on students' academic achievement and interest in the course. For the analysis of quantitative data, descriptive and inferential statistical methods were applied. Within the scope of descriptive statistics, the mean, standard deviation, skewness, and kurtosis values of the students' scores were calculated, and group differences were compared. Subsequently, in the inferential statistics phase, a One-Way Multivariate Analysis of Covariance (MANCOVA) was conducted to examine the effect of BilgeCan Bot on students' academic achievement and interest in science. MANCOVA allows testing the statistically significant effect of an independent variable on two or more dependent variables [103–105]. In this study, the post-test scores of the Academic Achievement Test (AAT<sub>2</sub>) and the Scale of Interest in Science Lesson (SISL<sub>2</sub>) were defined as dependent variables. During the analysis, the pre-test scores of the Scale of Interest in Science Lesson (SISL<sub>1</sub>) were controlled as a covariate to more precisely assess the effect of the independent variable on the dependent variables.

### 5.5.2. Analysis of Qualitative Data

In this study, students' views on chatbot use in science education were examined. Qualitative data were gathered through semi-structured interviews with experimental group students, BilgeCan Bot conversation records, and control group teacher interviews. Student interviews were analyzed through content analysis [106], with students anonymized as S1, S2, etc. BilgeCan Bot records underwent descriptive and thematic analyses. Variables such as Interaction Duration (ID), Number of Questions to the Bot (NQB), and Number of Dialogues with the Live Teacher (NDLT) were calculated based on estimated interaction times. Dialogue content was further examined using TF-IDF and n-gram techniques to extract key concepts. Student questions were classified according to Bloom's cognitive taxonomy [107]. Control group difficulties were thematically analyzed based on teacher interviews. Two researchers independently coded all qualitative data. Initial codes were generated and iteratively refined through discussion, resulting in broader themes. To ensure the reliability of coding, inter-coder agreement was calculated using Cohen's Kappa coefficient ( $\kappa = 0.84$ ), indicating substantial agreement.

## VI. Results

### 6.1. Descriptive Statistics

The descriptive statistics for the AAT<sub>1</sub>, AAT<sub>2</sub>, SISL<sub>1</sub>, and SISL<sub>2</sub> test results of the experimental and control groups are presented in Table 2.

**Table 2. Descriptive Statistics Results of the Groups for The Tests**

	Experimental group				Control group			
	AAT <sub>1</sub>	AAT <sub>2</sub>	SISL <sub>1</sub>	SISL <sub>2</sub>	AAT <sub>1</sub>	AAT <sub>2</sub>	SISL <sub>1</sub>	SISL <sub>2</sub>
N	26	26	26	26	26	26	26	26
Mean	12.08	20.12	98.62	106.62	11.27	17.27	96.62	94.04
SD	1.47	2.88	8.59	13.40	2.05	2.66	14.73	12.40
Skewness	-.14	.25	.18	-.70	.36	.14	-.14	.03
Kurtosis	-.42	-.84	-.44	-.02	.39	-1.16	-1.05	-1.03

The descriptive statistics for the AAT<sub>1</sub>, AAT<sub>2</sub>, SISL<sub>1</sub>, and SISL<sub>2</sub> test results of the experimental and control groups are presented in Table 2. AAT scores were evaluated out of 27 points, with each correct answer awarded 1 point. The experimental group's mean AAT<sub>1</sub> score was 12.08, and their mean AAT<sub>2</sub> score was 20.12. In the control group, these values were 11.27 and 17.27, respectively. Accordingly, an 8-point increase was observed in the experimental group, while a 6-point increase was recorded in the control group. SISL scores were evaluated out of 135 points. The experimental group's mean SISL score increased from 98.62 to 106.62, whereas the control group's mean score decreased from 96.62 to 94.04. These findings indicate that the experimental group showed greater improvement in both academic achievement and interest in science compared to the control group. Additionally, the skewness and kurtosis values for AAT<sub>1</sub>, AAT<sub>2</sub>, SISL<sub>1</sub>, and SISL<sub>2</sub> variables were within the range of -1.5 to +1.5, suggesting that the data were normally distributed [108, 109].

### 6.2. Inferential statistics

6.2.1 RQ1: Is There a Significant Difference in The Post-Test Scores of Academic Achievement and Science Interest Between Students Who Used Bilgecan Bot (Experimental Group) and Those Who Used The Textbook (Control Group), After Controlling for Pre-Test Scores?

The effects of BilgeCan Bot on students' academic achievement (AAT<sub>2</sub>) and science interest (SISL<sub>2</sub>) were analyzed using MANCOVA and ANCOVA. AAT<sub>1</sub> and SISL<sub>1</sub> were initially considered potential covariates, and their correlations with the dependent variables were tested. AAT<sub>1</sub> showed no significant correlation with AAT<sub>2</sub> or SISL<sub>2</sub> ( $p > .01$ ), whereas SISL<sub>1</sub> was significantly correlated with AAT<sub>2</sub> ( $r = .035$ ,  $p < .01$ ), suggesting its suitability as a covariate. Thus, SISL<sub>1</sub> was included in further analyses, while AAT<sub>1</sub> was excluded. MANCOVA assumptions—normality, homogeneity of variance, regression slope

homogeneity, multicollinearity, and independence—were tested and met. The analysis showed a significant effect of group on the combined dependent variables (Wilks' Lambda = .66,  $F(2, 48) = 12.04$ ,  $p < .05$ ,  $\eta^2 = .34$ , observed power = .99), indicating that BilgeCan Bot use explained 34% of the variance in AAT<sub>2</sub> and SISL<sub>2</sub>. To explore effects on each variable, separate ANCOVAs were conducted. The results are presented in Table 3.

**Table 3. ANCOVA Results**

Source	DV	Type III sum of squares	df	Mean square	F	Sig	Partial Eta squared	Observed power
Treatment	AAT <sub>2</sub>	92,75	1	92,75	13.60	.00	.22	.96
	SISL <sub>2</sub>	2194,20	1	2194,20	13.53	.00	.22	.96

Separate ANCOVA results showed significant differences in favor of the experimental group for both academic achievement (AAT<sub>2</sub>:  $F(1, 48) = 13.60$ ,  $p < .05$ ,  $\eta^2 = .22$ , observed power = .96) and science interest (SISL<sub>2</sub>:  $F(1, 48) = 13.53$ ,  $p < .05$ ,  $\eta^2 = .22$ , observed power = .96), indicating both theoretical and practical significance [110].

Adjusted means revealed that the experimental group outperformed the control group in academic achievement ( $M = 20.12$  vs.  $17.26$ ,  $SE = .52$ ) and science interest ( $M = 106.81$  vs.  $93.84$ ,  $SE = 2.57$ ). Pairwise comparisons confirmed that these differences were statistically significant for AAT<sub>2</sub> ( $MD = 2.86$ ,  $SE = .75$ ,  $p < .05$ , 95% CI [1.36, 4.36]) and for SISL<sub>2</sub> ( $MD = 12.97$ ,  $SE = 3.69$ ,  $p < .05$ , 95% CI [5.56, 20.39]).

Taken together, these findings indicate that the use of BilgeCan Bot significantly enhanced students' academic achievement and their interest in science, and that this effect was in favor of the experimental group.

### 6.2.2 RQ2: What are the Experimental Group Students' Opinions About Bilgecan Bot After the Intervention?

After the implementation, semi-structured interviews (SSIF) were conducted to gather the experimental group students' opinions about BilgeCan Bot. Responses were analyzed using content analysis, with similar expressions coded and grouped into themes, categories, and codes, supported by frequency values. Six of the 13 SSIF questions were associated with the theme "Using Chatbots for Learning," and the related findings are summarized in Table 4 (see Appendix D).

According to Table 4, all students reported that they used a chatbot for the first time, felt comfortable during the interaction, and found BilgeCan Bot easy to use. Sample student responses within this category are as follows:

*"I felt comfortable while using BilgeCan Bot and asked my questions without hesitation." (S12)*

*"BilgeCan answered my questions accurately and in detail. I liked that the responses came immediately." (S3)*

*"BilgeCan Bot provided information on different topics; it was helpful, and I was able to use it very easily." (S9)*

*"It was simple to use, and I could easily enter the questions I wanted to ask, and the bot answered them correctly." (S5)*

According to Table 4, most students reported encountering no problems while using the chatbot, receiving quick responses, finding it enjoyable and friendly, and expressing willingness to use it for other science topics and subjects. Example student responses for this category are presented below:

*"No, I did not encounter any problems." (S5)*

*"I liked the quick responses of BilgeCan Bot." (S18)*

*"Thanks to BilgeCan Bot, I find the science course more interesting and important." (Ö24)*

*"It was great that the active teachers sometimes responded to us." (S4)*

*"I wish all subjects could be taught with BilgeCan Bot." (S6)*

Students' criticisms of the chatbot mainly concerned difficulties in connecting with active teachers and the limited scope of available information. Example student responses for this category are presented below:

*"It doesn't answer some of my questions and cannot connect me to an active teacher." (S1)*

*"Sometimes when I log in, it says that there is no available teacher." (S19)*

The last seven questions in the SSIF belong to the theme "Benefits of Chatbots in Learning," and the findings related to this theme are presented in Table 5 (see Appendix E).

According to Table 5, most students reported that BilgeCan Bot supported their learning, improved their inquiry and thinking skills, made science lessons more engaging, and expressed a desire to use it for other topics. Some students also mentioned forming a sense of "small friendships" with the bot during solo study. Several students stated that they could understand complex topics more easily through the bot's simple and clear responses. For instance:

*"Thanks to BilgeCan Bot, I learned more about the Sun, Earth, and Moon, and understood the topic better." (S7)*

Others emphasized how the bot encouraged question-asking behaviors and helped improve their critical thinking:

*"Yes, it helped. Thanks to BilgeCan Bot, I learned how to ask questions about different concepts." (S23)*

*"As I kept asking questions to BilgeCan Bot, I felt that I started thinking differently and asking better questions." (S19)*

The chatbot's role in increasing students' interest and motivation was also highlighted:

*"BilgeCan Bot made the science lesson more fun and sparked more curiosity about the subject." (S6)*

*"When I was stuck at home, I used BilgeCan Bot. It made me feel like I wasn't alone and helped me continue studying."*

(S15)

*“I wish there was a BilgeCan Bot for the living things unit as well.” (S3)*

In addition, some students reflected on the lack of emotional connection, acknowledging that the bot had limitations in forming human-like bonds:

*“BilgeCan Bot cannot make you feel like a real friend because it cannot form an emotional bond.” (S9)*

*“Sometimes I wanted it to say something like ‘well done,’ but it just gave the answer and moved on.” (S11)*

Finally, a few students also commented on technical or contextual limitations, such as the bot's inability to understand complex or ambiguous questions:

*“Sometimes it didn't understand what I meant. I wish I could ask the teacher at those moments.” (S1)*

These statements indicate that the students were generally positive about BilgeCan Bot and found it useful as an academic and motivational support tool in science lessons, but were also aware of emotional and contextual limitations. Overall, the fact that students wanted to use the bot in different science subjects reveals that the overall satisfaction level was high.

### 6.2.3 RQ3: What is the Content of the Conversations that the Experimental Group Students Had with Bilgecan Bot?

The experimental group's interaction frequency and content usage with BilgeCan Bot were analyzed through descriptive and content analysis. Students typically used the bot twice a week during out-of-school hours for homework. Figure 11 shows weekly usage: 11 students in week one, rising to 18 in week two, dropping to 14 in week three, increasing to 16 in week four, peaking at 22 in week five, and falling to 15 in week six. These findings indicate that chatbot usage fluctuated weekly.

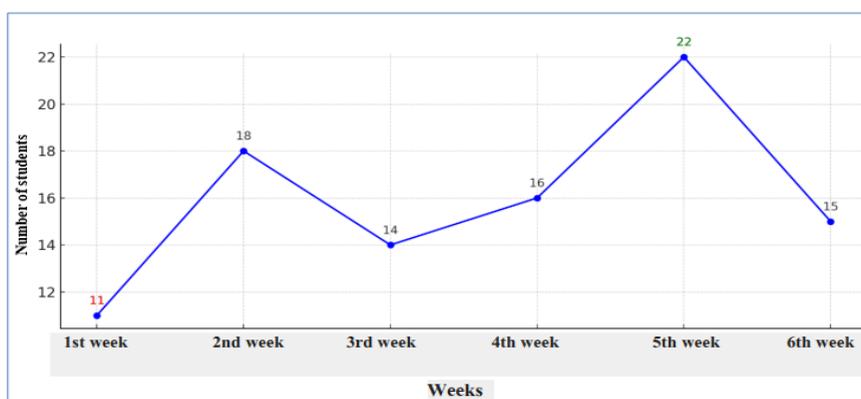


Figure 11. Number of students using bilgecan bot by week

As shown in Figure 12, the amount of time students spent with BilgeCan Bot varied across the weeks.

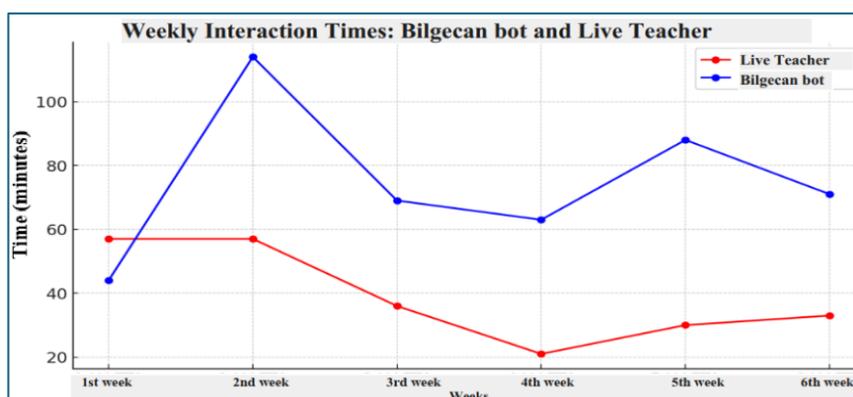


Figure 12. Interaction times per week with bilgecan bot and live tutor

Interaction time with BilgeCan Bot was lowest in the first week, peaked in the second, declined in the third and fourth, rose in the fifth, and dropped again in the sixth week. Interaction with live teachers remained stable initially, decreased mid-study, and increased in the final two weeks. These patterns indicate weekly shifts in students' preferences, with a notable rise in chatbot use during the fifth week.

Weekly interaction data were analyzed through variables such as Interaction Duration (ID), Number of Dialogues with BilgeCan Bot (NDBB), Number of Questions per Dialogue (NQDBB), Number of Questions Asked (NQABB), and Number of Dialogues with the Live Teacher (NDLT). S7 and S10 were the most active users, while S5 and S21 showed lower

interaction. Changes in interaction time corresponded with fluctuations in the number of questions asked. Figure 13 displays the weekly comparison of total questions directed to BilgeCan Bot and the live teacher.

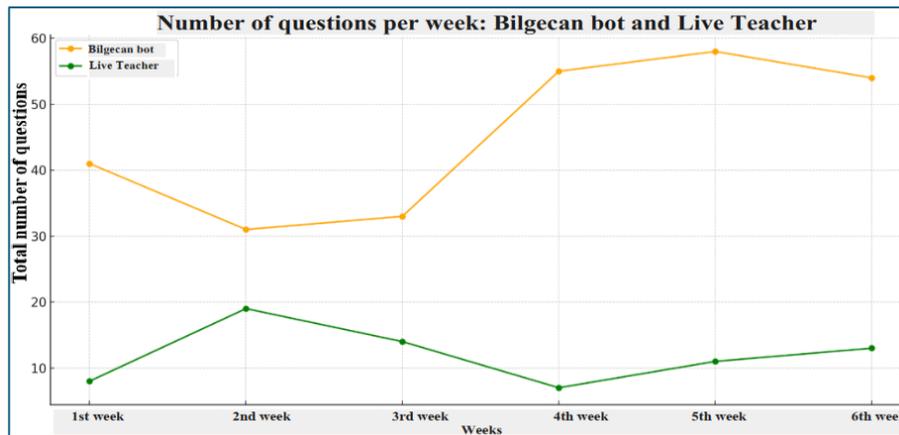


Figure 13. Weekly number of questions asked to BilgeCan bot and live teacher

When examining Fig. 13, it is evident that the number of questions directed to BilgeCan Bot was consistently higher each week than those directed to the live teacher, with a noticeable peak during the fourth and fifth weeks.

To further explore the content of students' interactions with BilgeCan Bot, a TF-IDF (Term Frequency–Inverse Document Frequency) analysis was conducted. Table 6 presents the top 10 keywords with the highest TF-IDF values identified across the six-week period (see Appendix F). According to Table 6, "Sun" dominated the first and second weeks, "Moon" became prominent in the third week, while "Moon's," "Moon," and "appears" were key in the fourth and fifth weeks. In the sixth week, keywords such as "Earth's," "time," and "rotation" stood out.

Additionally, conversation records were categorized according to Bloom's taxonomy (knowledge, comprehension, application, and analysis), and the results are presented in Table 7 (see Appendix G). The analysis showed that while knowledge-level questions increased over time, higher-order questions (comprehension, application, analysis) fluctuated, with notable peaks in certain weeks.

#### 6.2.4 RQ4: What Difficulties did the Control Group Students Encounter While Completing Their Homework Using the Textbook?

Difficulties experienced by control group students during homework completion were identified through post-lesson teacher interviews. Students reported problems such as conceptual misunderstandings, lack of examples, and insufficient guidance. They found textbook explanations inadequate for topics like the Sun's geometric structure, rotational movement, size comparisons between the Sun and Earth, and the Moon's surface features. The phases of the Moon were especially confusing, with students expressing a need for step-by-step support. Similar issues arose in modeling activities due to the absence of guidance. Students also indicated difficulty in understanding textbook content and noted that their questions remained unanswered. These findings suggest that textbooks alone were insufficient for supporting the comprehension of abstract concepts. The results are presented in Table 8 (see Appendix H).

Table 8 shows that students in the control group experienced two main difficulties while completing their homework: challenges in visualization and modeling, and a lack of interaction and support. Students reported struggling to understand abstract concepts and expressed that they did not receive sufficient guidance. These findings highlight the need for more visual and interactive support in instructional materials.

## VII. Discussion and Conclusion

This study investigated the effects of BilgeCan Bot, a Non-AI rule-based (NARB) chatbot designed to support science homework, on students' academic achievement and interest in science. It also explored experimental group students' perceptions of chatbot use and analyzed challenges faced by students relying solely on textbooks. Descriptive analyses revealed that students using BilgeCan Bot showed significantly greater improvements in both academic achievement and science interest compared to those using textbooks. MANCOVA results demonstrated that BilgeCan Bot significantly enhanced students' combined achievement and interest scores after controlling for prior interest, and ANCOVA results confirmed higher individual scores in the experimental group. Findings were interpreted within the framework of Cognitive Load Theory (CLT). Given the scarcity of studies focusing specifically on cognitive load in chatbot-based learning [111, 112], this research addresses an important gap. Although most prior work centers on AI-driven pedagogical agents [51, 113, 114], BilgeCan Bot—

despite lacking AI—meets the functional definition of a pedagogical agent by guiding cognitive processes and structuring homework tasks [115].

The findings indicated that BilgeCan Bot's simple, cognitively streamlined structure contributed positively to learning. Quantitative data showed that students using BilgeCan Bot achieved higher scores than the control group. This result aligns with previous research highlighting that structured digital guidance facilitates learning [51, 113]. Pedagogical agents enhance content interaction and academic performance [116, 117]. Wu et al. (2024) similarly demonstrated that guided learning enriched by student participation improves academic achievement [52]. Thus, BilgeCan Bot optimized students' cognitive resources, leading to achievement gains.

The study also found a statistically significant increase in students' interest in science. Qualitative data supported this, with students expressing that science lessons became more enjoyable through the bot. For example, one student remarked, "Science lessons were more fun thanks to the bot" (S6). This suggests that the bot fostered positive emotions toward learning. Bowman [118] and Dinçer & Doğanay [119] emphasized that personal, clear language and sustained attention through pedagogical agents are key factors in maintaining student interest. BilgeCan Bot's communication style likely contributed to these effects.

Regarding homework completion, students completed their homework more independently: "Thanks to the bot, I completed my homework without asking the teacher" (S3). This finding aligns with studies showing that interactive systems enhance engagement and boost homework motivation by reducing feelings of isolation [120, 52, 121]. Structured access to information helped students control their learning, thereby reducing extraneous cognitive load.

BilgeCan Bot's rule-based design minimized distractions by providing only essential information. Sakellariou et al. noted that highly complex AI systems can overwhelm students, whereas simple, goal-oriented designs support cognitive efficiency [122]. Chen et al. [117] and Mayer & Fiorella [123] also emphasized that overly interactive systems may increase cognitive burden in younger learners, favoring simpler designs.

Qualitative feedback indicated that the bot promoted active cognitive engagement: "As I asked questions, I started thinking about new things" (S23). This aligns with Sweller's [57] and Mayer & Moreno's [124] arguments that environments fostering meaningful cognitive processing are crucial for deep learning. Likewise, Ba et al. [125] and Bowman et al. [118] stressed the importance of pedagogical agents in supporting conceptual structure building. By encouraging inquiry, BilgeCan Bot fostered active learning engagement.

However, some students reported that the bot lacked emotional connection: "It's not like a real friend" (S9). This reflects findings by Wang & Lo [116] and Brandt et al. [126], emphasizing that chatbots may struggle to meet adolescents' expectations for emotional interaction. Therefore, future designs may benefit from incorporating emotional responsiveness to enhance pedagogical effectiveness.

In conclusion, this study demonstrated that BilgeCan Bot, despite operating as a non-AI rule-based system, effectively supported student learning by managing cognitive load, fostering engagement, and improving academic achievement and science interest. Within the framework of CLT, BilgeCan Bot reduced extraneous cognitive load by delivering essential information, balanced intrinsic load through structured explanations, and promoted germane load by encouraging inquiry-based learning. These outcomes indicate that simple, structured chatbots can significantly enhance student learning without relying on complex AI systems, especially in abstract domains like astronomy. Future research could examine the long-term effects of rule-based chatbot support on students' retention of science concepts, higher-order thinking skills, and intrinsic motivation, providing a broader understanding of its sustained educational impacts. From a policy perspective, it is recommended that national and regional education authorities pilot the use of rule-based chatbot systems, such as BilgeCan Bot, within science curricula to support homework completion, reduce cognitive overload, and enhance equitable access to structured academic guidance.

## VIII. Limitations

Although this study offers valuable insights, it has limitations. The sample was limited to 52 students from a single school, reducing generalizability. The chatbot was applied only to the "Sun, Earth, and Moon" unit. As a non-AI system, it lacked flexibility and could not track interaction time. The seven-week duration restricted assessment of long-term outcomes. Additionally, variations in students' internet access may have affected the level of interaction.

## IX. Recommendations

The findings suggest that non-AI rule-based chatbots like BilgeCan Bot can effectively support structured homework in science education. Future studies should explore their application in other subjects and across different age groups. Comparative research with AI-based systems is recommended within the Cognitive Load Theory framework. Integrating multimedia features may enhance user experience. Expanding chatbot functionality to involve teachers and parents could increase impact. Longitudinal studies are needed to assess effects on retention and higher-order thinking. Additionally, in-service training should be provided to help teachers integrate chatbots effectively into instructional practices.

### Acknowledgments

This study was adapted from the first author's PhD thesis.

### Authors' Contributions

TY conducted this study under the supervision of İD as part of a doctoral dissertation. TY designed the study, collected and analyzed the data, and drafted the manuscript in consultation with İD. Both authors read and approved the final manuscript.

### Competing Interests

The authors declare that they have no conflict of interest.

### References

- [1] L. Wang, Z. Liu, A. Liu, and F. Tao, "Artificial intelligence in product lifecycle management," *Int. J. Adv. Manuf. Technol.*, vol. 114, no. 3–4, pp. 771–796, 2021, doi: 10.1007/s00170-021-06882-1.
- [2] S. S. Barsoum, M. S. Elnagar, and B. M. Awad, "The effectiveness of using a cognitive style-based chatbot in developing science concepts and critical thinking skills among preparatory school pupils," *Eur. Sci. J.*, vol. 18, no. 22, p. 52, 2022, doi: 10.19044/esj.2022.v18n22p52.
- [3] R. Wu and Z. Yu, "Do AI chatbots improve students' learning outcomes? Evidence from a meta-analysis," *Br. J. Educ. Technol.*, vol. 55, no. 1, pp. 10–33, 2023, doi: 10.1111/bjet.13334.
- [4] Y. Liang, S. Ma, and C. Lin, "Chatbotification for web information systems: A pattern-based approach," in *Proc. 2022 IEEE 46th Annu. Comput., Softw. Appl. Conf. (COMPSAC)*, pp. 2290–2295, 2024, doi: 10.1109/compsac61105.2024.00368.
- [5] J. Q. Pérez, T. Daradoumis, and J. M. M. Puig, "Rediscovering the use of chatbots in education: A systematic literature review," *Comput. Appl. Eng. Educ.*, vol. 28, no. 6, pp. 1549–1565, 2020, doi: 10.1002/cae.22326.
- [6] N. L. Schroeder, O. O. Adesope, and R. B. Gilbert, "How effective are pedagogical agents for learning? A meta-analytic review," *J. Educ. Comput. Res.*, vol. 49, no. 1, pp. 1–39, 2013, doi: 10.2190/ec.49.1.a.
- [7] J. Schmidhuber, S. Schlögl, and C. Ploder, "Cognitive load and productivity implications in human-chatbot interaction," in *Proc. 2021 IEEE 2nd Int. Conf. Human-Machine Syst. (ICHMS)*, pp. 1–6, 2021, doi: 10.1109/ICHMS53169.2021.9582445.
- [8] X. Deng and Z. Yu, "A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education," *Sustainability*, vol. 15, no. 4, p. 2940, 2023, doi: 10.3390/su15042940.
- [9] F. R. Baskara, "Chatbots and flipped learning: Enhancing student engagement and learning outcomes through personalised support and collaboration," *Int. J. Recent Educ. Res.*, vol. 4, no. 2, pp. 223–238, 2023, doi: 10.46245/ijorer.v4i2.331.
- [10] F. Almasri, "Exploring the impact of artificial intelligence in teaching and learning of science: A systematic review of empirical research," *Res. Sci. Educ.*, vol. 54, no. 5, pp. 977–997, 2024, doi: 10.1007/s11165-024-10176-3.
- [11] V. Zenuni-Idrizi and L. Lutfiu-Kadriu, "The impact of regularly checking homework on students' success in primary school," *Univ. Teach. Educ. J.*, vol. 6, no. 11–12, pp. 178–183, 2024, doi: 10.62792/ut.education.v6.i11-12.p2653.
- [12] E. N. Mutayomba, "The impact of homework on teaching and learning in secondary schools: Case of Adventist secondary schools in Goma/D.R. Congo," *Int. J. Soc. Sci. Humanit. Invention*, vol. 10, no. 3, pp. 7823–7831, 2023, doi: 10.18535/ijsshi/v10i03.10.
- [13] S. B. Payne and E. Swanson, "Supporting families to motivate their middle school student during homework time," *Teach. Except. Child.*, vol. 55, no. 6, pp. 422–430, 2022, doi: 10.1177/00400599221099656.
- [14] L. Corno, "Homework is a complicated thing," *Educ. Res.*, vol. 25, no. 8, pp. 27–30, 1996, doi: 10.3102/0013189x025008027.
- [15] J. S. Stephen, "Identifying support systems and resources," in *Springer Texts in Education*, pp. 45–57, 2024, doi: 10.1007/978-3-031-54439-2\_4.
- [16] J. Lee, T. An, H. Chu, H. Hong, and S. N. Martin, "Improving science conceptual understanding and attitudes in elementary science classes through the development and application of a rule-based AI chatbot," *Asia-Pac. Sci. Educ.*, vol. 9, no. 2, pp. 365–412, 2023, doi: 10.1163/23641177-bja10070.
- [17] M. Taşcan, "The effect of science activities developed on astronomy education on fifth-grade students' spatial skills and academic achievement," Ph.D. dissertation, İnönü Univ., 2019.
- [18] F. Kara and N. Kefeli, "Perceptions of fifth-grade secondary school students towards the Sun, Earth, and Moon," *Necatibey Fac. Educ. Electron. J. Sci. Math. Educ.*, vol. 13, no. 2, pp. 1000–1014, 2019, doi: 10.17522/balikesirnef.560113.
- [19] N. Demir, "The effect of the 5E learning model on seventh-grade students' academic achievement and attitudes in teaching astronomy topics," Ph.D. dissertation, Erciyes Univ., 2020.
- [20] J. Xu, "Why do students have difficulties completing homework? The need for homework management," *J. Educ. Train. Stud.*, vol. 1, no. 1, 2013, doi: 10.11114/jets.v1i1.78.
- [21] A. Dhanani, "Using search engines in the teaching and learning process," *PEOPLE: Int. J. Soc. Sci.*, vol. 6, no. 1, pp. 682–689, 2020, doi: 10.20319/PIJSS.2020.61.682689.

- [22] S. Abbasi, H. Kazi, and N. N. Hussaini, "Effect of chatbot systems on students' learning outcomes," *Sylwan*, vol. 163, no. 10, pp. 49–63, 2019. [Online]. Available: [https://www.researchgate.net/publication/336373880\\_Effect\\_of\\_Chatbot\\_Systems\\_on\\_Student's\\_Learning\\_Outcomes](https://www.researchgate.net/publication/336373880_Effect_of_Chatbot_Systems_on_Student's_Learning_Outcomes)
- [23] B. A. Shawar and E. Atwell, "Chatbots: Are they really useful?" *J. Lang. Technol. Comput. Linguist.*, vol. 22, no. 1, pp. 29–49, 2007. [Online]. Available: <https://jcl.org/article/download/88/86>
- [24] S. S. Shaffie, M. F. Ouardin, S. R. Ismail, A. A. Rahman, I. M. Ibrahim, and A. F. Najmuddin, "Attending teaching and learning with chatbot," *Int. J. Acad. Res. Prog. Educ. Dev.*, vol. 13, no. 1, 2024, doi: 10.6007/ijarped/v13-i1/21107.
- [25] C. W. Okonkwo and A. Ade-Ibijola, "Chatbots applications in education: A systematic review," *Comput. Educ.: Artif. Intell.*, vol. 2, p. 100033, 2021, doi: 10.1016/j.caeai.2021.100033.
- [26] H.-A. Bekkar and Y. Chtouki, "Chatbots in education: A systematic literature review," in *Proc. IEEE Int. Conf. Syst., Commun. Coding*, pp. 637–644, 2024, doi: 10.1109/icscc62041.2024.10690334.
- [27] E. Loos, L. Ivan, and D. Leu, "'Save the Pacific Northwest tree octopus': a hoax revisited. Or: How vulnerable are school children to fake news?" *Inf. Learn. Sci.*, vol. 119, no. 9/10, pp. 514–528, 2018.
- [28] B. Lund, W. Ting, N. R. Mannuru, B. Nie, S. Shimray, and Z. Wang, "ChatGPT and a new academic reality: AI-written research papers and the ethics of the large language models in scholarly publishing," *SSRN Electron. J.*, 2023, doi: 10.2139/ssrn.4389887.
- [29] E. Feigerlova, H. Hani, and E. Hothersall-Davies, "A systematic review of the impact of artificial intelligence on educational outcomes in health professions education," *BMC Med. Educ.*, vol. 25, no. 1, 2025, doi: 10.1186/s12909-025-06719-5.
- [30] C. Papakostas, C. Troussas, A. Krouska, and C. Sgouropoulou, "A rule-based chatbot offering personalized guidance in computer programming education," in *Lect. Notes Comput. Sci.*, pp. 253–264, 2024, doi: 10.1007/978-3-031-63031-6\_22.
- [31] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cogn. Sci.*, vol. 12, no. 2, pp. 257–285, 1988, doi: 10.1016/0364-0213(88)90023-7.
- [32] F. Paas, A. Renkl, and J. Sweller, "Cognitive load theory and instructional design: Recent developments," *Educ. Psychol.*, vol. 38, no. 1, pp. 1–4, 2003, doi: 10.1207/s15326985ep3801\_1.
- [33] J. Weizenbaum, "ELIZA—A computer program for the study of natural language communication between man and machine," 1966. [Online]. Available: <https://web.stanford.edu/class/cs124/p36-weizenbaum.pdf>
- [34] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed. Pearson Education, Inc., 1995.
- [35] J. Yin, T. Goh, and Y. Hu, "Interactions with educational chatbots: The impact of induced emotions and students' learning motivation," *Int. J. Educ. Technol. High. Educ.*, vol. 21, no. 1, 2024, doi: 10.1186/s41239-024-00480-3.
- [36] H. Güldal and E. O. Dinçer, "Can rule-based educational chatbots be an acceptable alternative for students in higher education?" *Educ. Inf. Technol.*, 2024, doi: 10.1007/s10639-024-12977-5.
- [37] M. Černý, "Educational psychology aspects of learning with chatbots without artificial intelligence: Suggestions for designers," *Eur. J. Invest. Health Psychol. Educ.*, vol. 13, no. 2, pp. 284–305, 2023, doi: 10.3390/ejihpe13020022.
- [38] O. V. Deryugina, "Chatterbots," *Sci. Tech. Inf. Process.*, vol. 37, no. 2, pp. 143–147, 2010, doi: 10.3103/s0147688210020097.
- [39] S. Roos, "Chatbots in education: A passing trend or a valuable pedagogical tool?" M.S. thesis, 2018. [Online]. Available: <https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-355054>
- [40] Y. V. Lakshmi and I. Majid, "Chatbots in the education system," *Univ. News: A Wkly. J. High. Educ.*, vol. 60, no. 8, pp. 15–18, 2022. [Online]. Available: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4039535](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4039535)
- [41] İ. Özkoç, K. Doğan, and G. Köseali, "Artificial intelligence supported chatbot usage in ERMS applications," 2019. [Online]. Available: [https://www.researchgate.net/publication/338336314\\_EBYS\\_Uygulamalarinda\\_Yapay\\_Zeka\\_Destekli\\_Chatbot\\_Sohbet\\_Robotu\\_Kullanimi\\_-\\_Artificial\\_Intelligence\\_Supported\\_Chatbot\\_Usage\\_in\\_ERMS\\_Applications](https://www.researchgate.net/publication/338336314_EBYS_Uygulamalarinda_Yapay_Zeka_Destekli_Chatbot_Sohbet_Robotu_Kullanimi_-_Artificial_Intelligence_Supported_Chatbot_Usage_in_ERMS_Applications)
- [42] L. C. Klopfenstein, S. Delpriori, S. Malatini, and A. Bogliolo, "The rise of bots: A survey of conversational interfaces, patterns, and paradigms," in *Proc. 2017 Conf. Designing Interactive Systems*, pp. 555–565, 2017. [Online]. Available: <https://dl.acm.org/doi/abs/10.1145/3064663.3064672>
- [43] S. Hussain, O. A. Sianaki, and N. Ababneh, "A survey on conversational agents/chatbots classification and design techniques," in *Adv. Intell. Syst. Comput.*, pp. 946–956, 2019, doi: 10.1007/978-3-030-15035-8\_93.
- [44] H. Chen, X. Liu, D. Yin, and J. Tang, "A survey on dialogue systems," *ACM SIGKDD Explor. Newsl.*, vol. 19, no. 2, pp. 25–35, 2017, doi: 10.1145/3166054.3166058.
- [45] E. Adamopoulou and L. Moussiades, "Chatbots: History, technology, and applications," *Mach. Learn. Appl.*, vol. 2, p. 100006, 2020, doi: 10.1016/j.mlwa.2020.100006.
- [46] D. Madhu, C. J. N. Jain, E. Sebastain, S. Shaji, and A. Ajayakumar, "A novel approach for medical assistance using trained chatbot," in *Proc. 2017 Int. Conf. Inventive Commun. Comput. Technol. (ICICCT)*, pp. 243–246, 2017, doi: 10.1109/ICICCT.2017.7975195.
- [47] K. F. Hew, W. Huang, J. Du, and C. Jia, "Using chatbots to support student goal setting and social presence in fully online activities: Learner engagement and perceptions," *J. Comput. High. Educ.*, vol. 35, no. 1, pp. 40–68, 2023, doi: 10.1007/s12528-022-09338-x.

- [48] S. A. Abdul-Kader and J. Woods, "Survey on chatbot design techniques in speech conversation systems," *Int. J. Adv. Comput. Sci. Appl.*, vol. 6, no. 7, 2015, doi: 10.14569/ijacsa.2015.060712.
- [49] A. Badii, A. Adetoye, D. Patel, and K. Hameed, "Efficient FPGA-Based Regular Expression Pattern Matching," 2015.
- [50] L. Bradeško and D. Mladenčić, "A survey of chatbot systems through a Loebner Prize competition," in *Proc. Slovenian Lang. Technol. Soc. 8th Conf. Lang. Technol.*, vol. 2, pp. 34–37, 2012. [Online]. Available: [https://nl.ijs.si/isjt12/proceedings/isjt2012\\_06.pdf](https://nl.ijs.si/isjt12/proceedings/isjt2012_06.pdf)
- [51] R. Moreno and R. Mayer, "Interactive multimodal learning environments," *Educ. Psychol. Rev.*, vol. 19, no. 3, pp. 309–326, 2007, doi: 10.1007/s10648-007-9047-2.
- [52] C. Wu, Y. Li, H. Yang, X. Wang, X. Li, and B. Jing, "The greater the interaction, the stronger the learning performance? Examining pedagogical agents' interactive presence in instructional videos," *Appl. Cogn. Psychol.*, vol. 38, no. 5, 2024, doi: 10.1002/acp.4242.
- [53] P. Ayres and F. Paas, "Cognitive load theory: New directions and challenges," *Appl. Cogn. Psychol.*, vol. 26, no. 6, pp. 827–832, 2012, doi: 10.1002/acp.2882.
- [54] R. Winkler and M. Söllner, "Unleashing the potential of chatbots in education: A state-of-the-art analysis," *Acad. Manag. Proc.*, vol. 2018, no. 1, p. 15903, 2018, doi: 10.5465/ambpp.2018.15903abstract.
- [55] P. A. Kirschner, J. Sweller, and R. E. Clark, "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching," *Educ. Psychol.*, vol. 41, no. 2, pp. 75–86, 2006, doi: 10.1207/s15326985ep4102\_1.
- [56] B. K. Mungai, P. K. K. Omieno, M. Egessa, and P. N. Manyara, "AI chatbots in LMS: A pedagogical review of cognitive, constructivist, and adaptive principles," *Eng. Technol. J.*, vol. 9, no. 8, 2024, doi: 10.47191/etj/v9i08.15.
- [57] J. Sweller, P. Ayres, and S. Kalyuga, *Cognitive Load Theory*. Springer New York, 2011, doi: 10.1007/978-1-4419-8126-4.
- [58] J. J. G. Van Merriënboer and J. Sweller, "Cognitive load theory and complex learning: Recent developments and future directions," *Educ. Psychol. Rev.*, vol. 17, no. 2, pp. 147–177, 2005, doi: 10.1007/s10648-005-3951-0.
- [59] S. Gumilar, "Conceptual understanding and visual representation: Are they still ordinary or extraordinary issues?" *Res. Phys. Educ.*, vol. 2, no. 2, pp. 1–2, 2023, doi: 10.31980/ripe.v2i2.29.
- [60] S. Mi, S. Lu, and H. Bi, "Trends and foundations in research on students' conceptual understanding in science education: A method based on the structural topic model," *J. Baltic Sci. Educ.*, vol. 19, no. 4, pp. 551–568, 2020, doi: 10.33225/JBSE/20.19.551.
- [61] X. Huang, "Application of artificial intelligence app in quality evaluation of primary school science education," *Educ. Stud.*, vol. 50, no. 6, pp. 1215–1235, 2022, doi: 10.1080/03055698.2022.2066462.
- [62] J. P. Adams and T. F. Slater, "Astronomy in the national science education standards," *J. Geosci. Educ.*, vol. 48, no. 1, pp. 39–45, 2000, doi: 10.5408/1089-9995-48.1.39.
- [63] P. Hubber, "Year 8 students' understanding of astronomy as a representational issue: Insights from a classroom video study," 2009. [Online]. Available: <http://dro.deakin.edu.au/eserv/DU:30021664/rentschler-year8students-2009.pdf>
- [64] P. M. Sadler, "Astronomy's conceptual hierarchy," in *Astronomy Education: Current Developments, Future Coordination* (ASP Conf. Ser., Vol. 89), J. R. Percy, Ed., p. 46, 1996.
- [65] E. Gazit, D. Chen, and Y. Yair, "Developing understanding of basic astronomical concepts by using a virtual solar system," in *Proc. Int. Conf. Learn. Sci.*, p. 601, 2004. [Online]. Available: <https://repository.isls.org/handle/1/4019>
- [66] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting with educational chatbots: A systematic review," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 973–1018, 2023, doi: 10.1007/s10639-022-11177-3.
- [67] S. Wollny, J. Schneider, D. Di Mitri, J. Weidlich, M. Rittberger, and H. Drachsler, "Are we there yet? A systematic literature review on chatbots in education," *Front. Artif. Intell.*, vol. 4, p. 654924, 2021, doi: 10.3389/frai.2021.654924.
- [68] P. Chandler and J. Sweller, "Cognitive load theory and the format of instruction," *Cogn. Instr.*, vol. 8, no. 4, pp. 293–332, 1991, doi: 10.1207/s1532690xci0804\_2.
- [69] N. D. Septiyanti, M. I. Luthfi, and D. Darmawansah, "Effect of chatbot-assisted learning on students' learning motivation and its pedagogical approaches," *Khazanah Informatika: J. Ilmu Komput. Informatika*, vol. 10, no. 1, pp. 69–77, 2024, doi: 10.23917/khif.v10i1.4246.
- [70] H. Cooper, "Synthesis of research on homework," *Educ. Leadersh.*, vol. 47, no. 3, pp. 85–91, 1989. [Online]. Available: [https://files.ased.org/staticfiles/ased/pdf/journals/ed\\_lead/el198911\\_cooper.pdf](https://files.ased.org/staticfiles/ased/pdf/journals/ed_lead/el198911_cooper.pdf)
- [71] U. Trautwein, A. Niggli, I. Schnyder, and O. Lüdtke, "Between-teacher differences in homework assignments and the development of students' homework effort, homework emotions, and achievement," *J. Educ. Psychol.*, vol. 101, no. 1, pp. 176–189, 2009, doi: 10.1037/0022-0663.101.1.176.
- [72] E. Güven Yıldırım and A. N. Köklükaya, "Determination of primary and middle school students' interest levels in science subjects," *Amasya Univ. J. Educ. Fac.*, vol. 5, no. 1, pp. 1–22, 2016. [Online]. Available: <https://dergipark.org.tr/en/pub/amauefd/issue/24319/257699>
- [73] R. Mintz, S. Litvak, and Y. Yair, "3D-virtual reality in science education: An implication for astronomy teaching," *J. Comput. Math. Sci. Teach.*, vol. 20, no. 3, pp. 293–305, 2001. [Online]. Available: <https://www.academia.edu/download/31690325/JCMST203293.pdf>

- [74] H. Cooper, J. C. Robinson, and E. A. Patall, "Does homework improve academic achievement? A synthesis of research, 1987–2003," *Rev. Educ. Res.*, vol. 76, no. 1, pp. 1–62, 2006, doi: 10.3102/00346543076001001.
- [75] S. Dettmers, U. Trautwein, and O. Lüdtke, "The relationship between homework time and achievement is not universal: Evidence from multilevel analyses in 40 countries," *Sch. Eff. Sch. Improv.*, vol. 20, no. 4, pp. 375–405, 2009, doi: 10.1080/09243450902904601.
- [76] J. Xu and H. Wu, "Self-regulation of homework behavior: Homework management at the secondary school level," *J. Educ. Res.*, vol. 106, no. 1, pp. 1–13, 2012, doi: 10.1080/00220671.2012.658457.
- [77] J. Hattie and H. Timperley, "The power of feedback," *Rev. Educ. Res.*, vol. 77, no. 1, pp. 81–112, 2007, doi: 10.3102/003465430298487.
- [78] D. C. Moos and A. Ringdal, "Self-regulated learning in the classroom: A literature review on the teacher's role," *Educ. Res. Int.*, vol. 2012, pp. 1–15, 2012, doi: 10.1155/2012/423284.
- [79] K. A. Renninger and S. Hidi, "Interest development and learning," in *Cambridge University Press eBooks*, pp. 265–290, 2019, doi: 10.1017/9781316823279.013.
- [80] A. Kuzmin, A. Dukhanov, and S. Kraev, "Approach to intelligent support in educational program design using X-matrices and a domain-specific language," in *Proc. 2022 Int. Russ. Autom. Conf. (RusAutoCon)*, pp. 920–925, 2024, doi: 10.1109/rusautocon61949.2024.10694424.
- [81] Y. Lin and J. Ye, "Development of an educational chatbot system for enhancing students' biology learning performance," *J. Internet Technol.*, vol. 24, no. 2, pp. 275–281, 2023, doi: 10.53106/160792642023032402006.
- [82] R. Chocarro, M. Cortiñas, and G. Marcos-Matás, "Teachers' attitudes towards chatbots in education: A technology acceptance model approach considering the effect of social language, bot proactiveness, and users' characteristics," *Educ. Stud.*, vol. 49, no. 2, pp. 295–313, 2021, doi: 10.1080/03055698.2020.1850426.
- [83] A. S. Al Darayseh, "Acceptance of artificial intelligence in teaching science: Science teachers' perspective," *Comput. Educ.: Artif. Intell.*, vol. 4, p. 100132, 2023, doi: 10.1016/j.caeai.2023.100132.
- [84] F. T. M. Ayasrah, K. Alarabi, and H. A. A. Fattah, "A secure technology environment and AI's effect on science teaching: Prospective science teachers," *Migration Lett.*, vol. 20, no. S2, pp. 289–302, 2023. [Online]. Available: <https://migrationletters.com/index.php/ml/article/view/3687>
- [85] B. Kohli, T. Choudhury, S. Sharma, and P. Kumar, "A platform for human-chatbot interaction using Python," *IEEE Xplore*, 2018, doi: 10.1109/ICGCIoT.2018.8753031.
- [86] O. A. Mehdi, H. Ibrahim, and L. S. Affendey, "Instance-based matching using regular expression," *Procedia Comput. Sci.*, vol. 10, pp. 688–695, 2012, doi: 10.1016/j.procs.2012.06.088.
- [87] M. McTear, Z. Callejas, and D. Griol, *The Conversational Interface*. Springer, 2016, doi: 10.1007/978-3-319-32967-3.
- [88] S. Toraman, "Mixed methods research: A brief history, definitions, perspectives, and key elements," *Qual. Soc. Sci.*, vol. 3, no. 1, pp. 1–29, 2021, doi: 10.47105/nsb.847688.
- [89] L. Cohen, L. Manion, and K. Morrison, *Research Methods in Education*, 6th ed. London: Routledge, 2007.
- [90] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 2nd ed. Thousand Oaks, CA: Sage, 2003.
- [91] Ş. Büyükoztürk, *Handbook of Data Analysis for Social Sciences: Statistics and Research Design*. Ankara: Pegem Academy, 2018.
- [92] J. M. Cortina, "What is coefficient alpha? An examination of theory and applications," *J. Appl. Psychol.*, vol. 78, no. 1, pp. 98–104, 1993, doi: 10.1037/0021-9010.78.1.98.
- [93] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed. New York: McGraw-Hill, 1994.
- [94] K. S. Taber, "The use of Cronbach's alpha when developing and reporting research instruments in science education," *Res. Sci. Educ.*, vol. 48, no. 6, pp. 1273–1296, 2018, doi: 10.1007/s11165-016-9602-2.
- [95] M. F. Turgut, *Measurement and Evaluation in Education*, 9th ed. Ankara: Saydam, 1992.
- [96] C. Laçın Şimşek and H. Nuhoglu, "Development of a valid and reliable interest scale towards science topics," *Sakarya Univ. J. Educ. Fac.*, no. 18, pp. 28–41, 2009. [Online]. Available: <https://dergipark.org.tr/en/pub/sakaefd/issue/11214/133927>
- [97] E. Ş. Karalar, "Investigation of middle school students' interest in science subjects in terms of various variables," M.S. thesis, Muğla Sıtkı Koçman Univ., 2018.
- [98] H. Kendirli, "The effect of Educational Informatics Network (EBA)-supported science applications on seventh-grade students' interest in science," M.S. thesis, Niğde Ömer Halisdemir Univ., 2017.
- [99] F. Şimşek, "The effect of STEM activities on students' attitudes, interest, scientific process skills in science and student opinions," *Turk. J. Comput. Math. Educ.*, vol. 10, no. 3, pp. 654–679, 2019, doi: 10.16949/turkbilmat.470261.
- [100] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed. Sage Publications, 1994.
- [101] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, p. 159, 1977, doi: 10.2307/2529310.
- [102] K. Krippendorff, "On the reliability of unitizing continuous data," *Sociol. Methodol.*, pp. 47–76, 1995.
- [103] J. Fraenkel, N. Wallen, and H. Hyun, *How to Design and Evaluate Research in Education*, 10th ed. New York: McGraw-Hill Education, 1993.

- [104] J. Pallant, *SPSS Survival Manual: A Step-by-Step Guide to Data Analysis Using IBM SPSS*, 6th ed. New York: Open University Press, 2016.
- [105] V. Sönmez and G. F. Alacapınar, *Illustrated Scientific Research Methods*. Ankara: Anı Publishing, 2019.
- [106] A. Yıldırım and H. Şimşek, *Qualitative Research Methods in the Social Sciences*, 5th ed. Ankara: Seçkin, 2006.
- [107] V. Clarke and V. Braun, "Thematic analysis," *J. Positive Psychol.*, vol. 12, no. 3, pp. 297–298, 2016, doi: 10.1080/17439760.2016.1262613.
- [108] Ş. Büyüköztürk, *Handbook of Data Analysis for Social Sciences: Statistics, Research Design, SPSS Applications, and Interpretation*. Ankara: Pegem Academy, 2014.
- [109] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics*, 7th ed. New York: Allyn & Bacon/Pearson Education, 2013.
- [110] J. T. E. Richardson, "Eta squared and partial eta squared as measures of effect size in educational research," *Educ. Res. Rev.*, vol. 6, no. 2, pp. 135–147, 2011, doi: 10.1016/j.edurev.2010.12.001.
- [111] J. Yin, T.-T. Goh, and Y. Hu, "Using a chatbot to provide formative feedback: A longitudinal study of intrinsic motivation, cognitive load, and learning performance," *IEEE Trans. Learn. Technol.*, vol. 17, pp. 1404–1415, 2024, doi: 10.1109/tlt.2024.3364015.
- [112] Q. N. Nguyen, A. Sidorova, and R. Torres, "User interactions with chatbot interfaces vs. menu-based interfaces: An empirical study," *Comput. Human Behav.*, vol. 128, p. 107093, 2021, doi: 10.1016/j.chb.2021.107093.
- [113] S. Park, "The effects of social cue principles on cognitive load, situational interest, motivation, and achievement in pedagogical agent multimedia learning," *Educ. Technol. Soc.*, vol. 18, no. 4, pp. 211–229, 2015.
- [114] N. L. Schroeder, "The influence of a pedagogical agent on learners' cognitive load," *Educ. Technol. Soc.*, vol. 20, no. 4, pp. 138–147, 2017.
- [115] A. L. Baylor, "Expanding pre-service teachers' metacognitive awareness of instructional planning through pedagogical agents," *Educ. Technol. Res. Dev.*, vol. 50, no. 2, pp. 5–22, 2002.
- [116] Y. Wang and C. Lo, "The effects of response time on older and young adults' interaction experience with chatbot," *BMC Psychol.*, vol. 13, no. 1, 2025, doi: 10.1186/s40359-025-02459-9.
- [117] B. Chen, G.-H. Hwang, and S.-H. Wang, "Gender differences in cognitive load when applying game-based learning with intelligent robots," *Educ. Technol. Soc.*, vol. 24, no. 3, pp. 102–115, 2021, doi: 10.30191/ETS.202107\_24(3).0008.
- [118] C. D. D. Bowman, "Student use of animated pedagogical agents in a middle school science inquiry program," *Br. J. Educ. Technol.*, vol. 43, no. 3, pp. 359–375, 2011, doi: 10.1111/j.1467-8535.2011.01198.x.
- [119] S. Dinçer and A. Doğanay, "The effects of multiple-pedagogical agents on learners' academic success, motivation, and cognitive load," *Comput. Educ.*, vol. 111, pp. 74–100, 2017, doi: 10.1016/j.compedu.2017.04.005.
- [120] M. A. S. N. Nunes, L. L. Dihl, L. M. Fraga, C. R. Woszezenki, L. Oliveira, D. J. Francisco, and M. Notargiacomo, "Animated pedagogical agent in the intelligent virtual teaching environment," *Interact. Educ. Multimedia*, vol. 4, pp. 53–60, 2002.
- [121] F. S. Al-Hafdi and S. M. AlNajdi, "The effectiveness of using chatbot-based environment on learning process, students' performances and perceptions: A mixed exploratory study," *Educ. Inf. Technol.*, vol. 29, no. 15, pp. 20633–20664, 2024, doi: 10.1007/s10639-024-12671-6.
- [122] S. Sakellariou, A. Molohidis, and E. Hatzikraniotis, "On the effectiveness of pedagogical agents," *Int. J. Teach. Learn. Sci. (IJTLS-102)*, 2024, doi: 10.47991/2024/IJTLS-102.
- [123] R. E. Mayer and L. Fiorella, *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, 2022.
- [124] R. E. Mayer and R. Moreno, "Nine ways to reduce cognitive load in multimedia learning," *Educ. Psychol.*, vol. 38, no. 1, pp. 43–52, 2003, doi: 10.1207/S15326985EP3801\_6.
- [125] S. Ba, D. Stein, Q. Liu, T. Long, K. Xie, and L. Wu, "Examining the effects of a pedagogical agent with dual-channel emotional cues on learner emotions, cognitive load, and knowledge transfer performance," *J. Educ. Comput. Res.*, vol. 59, no. 6, pp. 1114–1134, 2021, doi: 10.1177/0735633121992421.
- [126] Å. Brandt, M. P. Jensen, M. S. Søberg, S. D. Andersen, and T. Sund, "Information and communication technology-based assistive technology to compensate for impaired cognition in everyday life: A systematic review," *Disabil. Rehabil.: Assist. Technol.*, vol. 15, no. 7, pp. 810–824, 2020, doi: 10.1080/17483107.2020.1765032.

## APPENDIX

### Appendix A

#### Lesson Implementation Checklist

Number	Processes	Done	Not Done
1	Checking the classroom's physical environment (lighting, heating, arrangement) before lesson begins		
2	Checking the seating arrangement suitability for lesson delivery		
3	Ensuring classroom quietness by the teacher		
4	Conducting attendance and identifying absent students		
5	Establishing a free and open environment for student expression		
6	Conducting lessons in accordance with the planned curriculum		
7	Encouraging student participation during lessons		
8	Ensuring accessibility of lesson materials		
9	Providing equal opportunities for students to ask questions and participate		
10	Effectively using technological tools during the lesson		
11	Allocating time for students to summarize and reflect on learning at the end		
12	Allowing time for students to share and reflect their understandings		
13	Giving feedback opportunities related to learning processes		
14	Clearly communicating learning objectives at the beginning and end of lessons		

### Appendix B

#### Semi-Structured Interview Form (SSIF) Regarding the Use of Educational Chatbots (Ecs)

1. Have you ever used chatbot technology before?
2. Did you encounter any problems while using the chatbot (BilgeCan Bot)? Please provide examples.
3. Did you feel comfortable while asking questions to the chatbot (BilgeCan Bot)? Please explain your thoughts.
4. What features of the chatbot (BilgeCan Bot) did you like? Please provide examples.
5. What features of the chatbot (BilgeCan Bot) did you dislike? Please provide examples.
6. What features would you like the chatbot (BilgeCan Bot) to have? Please explain and provide examples.
7. Did using the chatbot (BilgeCan Bot) help you learn the topics in the unit "The Sun, Earth, and Moon"? Please explain your thoughts.
8. Did using the chatbot (BilgeCan Bot) contribute to improving your questioning skills? Please explain your thoughts.
9. Did using the chatbot (BilgeCan Bot) contribute to improving your critical thinking skills? Please explain your thoughts.
10. Did using the chatbot (BilgeCan Bot) change your perspective towards the Science course? Please explain your thoughts.
11. Would you like to use the chatbot (BilgeCan Bot) to learn other units in the Science course? Please explain why, providing examples.
12. Would you like to use the chatbot (BilgeCan Bot) in other subjects outside the Science course? Please provide examples.
13. Do you think the chatbot (BilgeCan Bot) could be a companion for you when you are alone at home? Please explain your thoughts.

Appendix C

Sample Dialogue Between Students and the Live Teachers

Student S5–Teacher Dialogue

Dialogues

Hello, teacher

Hello

Teacher, I would like to ask something

Of course, you can ask

Why is the Sun yellow?

The Sun's color is not actually yellow

The Sun is like the color of a fireball?

Yes, its color is neither exactly yellow nor exactly orange

I understand

I understand

Thank you

You're welcome

Student S12–Teacher Dialogue

Hello, teacher

Hello

Teacher, I want to learn what a Full Moon is

A Full Moon is the phase when the entire face of the Moon, as seen from Earth, is fully illuminated

In other words, during a Full Moon, no part of the Moon's surface facing the Earth remains in darkness, and the Moon appears very bright

I understand, teacher

Thank you

You're welcome

Do you have any other questions?

No, I don't

Okay, see you later

See you later, teacher

## Appendix D

Table 4. Thematic Findings of the First 6 Questions in the SSIF

Interview Question	Theme (Using chatbot to learn)	Category	Code	Frequency
1st	Experience History	Use First	No, this is my first time using it	26
2nd	Issues and Challenges	Technical Issues	Delayed Responses	4
		Content	No problem Missing Content	19 3
3rd	User Comfort	Comfort	Comfortable to Use	26
4th	Liked Features	Answer Quality	Accurate and Detailed Answers	7
		Speed and Instant Response	Quick Replies	11
		Ease of Use	Easy to use	26
		Active Teachers Menu	Like	5
5th	Disliked Features	Instructive and Entertaining	Cute & Instructive	16
		Answer Quality	Not Responding	3
		Active Teachers	Inability to connect with active teachers	8
6th	Desired Features	Limited Subject Areas	Restricted Areas of Information	12
		Broad Subject Areas	Use in Other Science Subjects	21
		Personalization	Use in Different Courses	15
		Fun & Games	Personalized Responses	5
		Dialogue	Useful Games	9
			Chat	16

1. Have you used chatbot technology before?"
2. Did you encounter any problems while using BilgeCan Bot?
3. Did you feel comfortable asking questions to BilgeCan Bot? Write down your thoughts
4. What features of BilgeCan Bot do you like? Give an example.
5. Which features of BilgeCan Bot did you not like? Give an example.
6. What other features would you like BilgeCanBot to have other than its current features? Give an example

Appendix E

**Table 5. Thematic Findings of the Last 7 Questions in the SSIF**

Interview Question	Theme (Benefits of Chatbot in Learning)	Category	Code	Frequency
7th	Benefits of Chatbot for Learning	Help to Learn	Provision of Useful Information	22
		Better Comprehension	Better Understanding of the Topic	18
		Motivation and Interest	Motivating and Engaging	13
8th	Development of Question Inquiry Skills	Contribution to Question Inquiry Skills	Contributed	17
			It contributed a little bit	7
			No, it never happened	2
9th	Development of Thinking Skills	Contribution to Thinking Skills	Contributed	23
			It wasn't too much	3
		Sense of Self-Confidence	Self-confidence	8
10th	Change in the Perspective of Science Courses	Interest in Science Course	Interest in Science Courses Is Increasing	18
		Improvement in Learning and Comprehension	Progress in Understanding Science Subjects	14
		Positive perception of the course	Enjoying Science Class	16
11th	Willingness to Learn Other Topics with Chatbot	Positive Attention	Interest and Willingness to Learn Other Units	26
		Fun and Intuitive Learning	Ease and Pleasure of Learning with Chatbot,	16
		Independence from Time and Space	Flexibility of Learning with Chatbot	13
12th	Willingness to Use the Chatbot in Other Courses	Getting General Information	To have knowledge in different courses	26
		Help with Exams and Assignments	Get Help with Exams and Assignments	8
		Practical and Fast Information Access	Fast Learning in Other Lessons	11
		Learning Without Teacher Assistance	Obtaining knowledge without the help of a teacher	16
13th	Making Friends with a Chatbot	Structure and Capabilities of the Chatbot	The Limits of the Chatbot	6
		Human-Chatbot Relationship	Chatbot's ability to make friends	14
		Ability to establish emotional connection	The Chatbot's Lack of Emotion	11

7. Was using BilgeCan Bot helpful in learning about "Sun, Earth and Moon"? Write down your thoughts.

8. Has using BilgeCan Bot helped you improve your questioning skills? Write down your thoughts.

9. Did using BilgeCan Bot help you improve your thinking skills? Write down your thoughts.

10. Did using BilgeCan Bot change your perspective on science class? Write down your thoughts.

11. Would you like to use BilgeCan Bot to learn other subjects in science class? Give an example and explain why.

12. Would you like to use BilgeCanBot in other courses besides science? "Give me an example."

13. Can using BilgeCan Bot accompany you when you are alone at home? Write down your thoughts.

## Appendix F

Table 6. Token and TF-IDF Values of Weekly Chats

1st week		2nd week		3rd week		4th week		5th week		6th week	
Token	TFxIDF	Token	TFxIDF	Token	TFxIDF	Token	TFxIDF	Token	TFxIDF	Token	TFxIDF
sun	0,853	sun	0,823	moon	0,589	moon	0,618	moon	0,682	moon	0,408
sun	0,265	sun	0,206	month	0,295	month	0,507	month	0,364	month	0,306
keeping	0,235	sometimes	0,171	surface	0,236	appears	0,265	keeping	0,341	sometimes	0,306
time	0,147	look	0,171	world	0,177	sometimes	0,221	sometimes	0,25	world	0,272
appears	0,118	stains	0,171	moon	0,177	time	0,221	sun	0,227	sun	0,238
down	0,088	same	0,137	craters	0,177	same	0,176	same	0,136	same	0,238
night	0,088	bright	0,137	occurred	0,177	phases	0,132	appears	0,136	time	0,204
spots	0,088	a lot	0,103	around	0,177	around	0,088	time	0,136	around	0,204
same	0,088	time	0,103	rotating	0,177	occurred	0,088	day	0,114	day	0,17
light	0,088	size	0,103	month	0,177	month	0,088	phases	0,091	return	0,17

## Appendix G

Table 7. Thematic Findings of the Chat Recordings Between Students and Bilgecan Bot

Week	Question level	Number of questions	Sample questions
1st	Knowledge	11	What are the layers of the Sun?
	Understand	11	How does a solar eclipse happen?
	Application	11	When does a solar eclipse happen and how can I watch it?
	Analysis	8	Why doesn't the sun always look like it's in the same place?
2nd	Knowledge	9	Why does the sun look yellow?
	Understand	8	Why should I wear special glasses while watching a solar eclipse?
	Application	8	How can we see sunspots?
	Analysis	6	Why are Galileo's observations about sunspots important?
3rd	Knowledge	10	Is there an atmosphere on the Moon?
	Understand	8	How does the Moon revolve around the Earth?
	Application	8	How did craters form?
	Analysis	7	Why does the moon sometimes appear completely round at night, sometimes in a half or crescent shape?
4th	Knowledge	14	How long does it take for the Moon to revolve around an Earth?
	Understand	13	How can the moon be so bright at night?
	Application	15	How to distinguish the phases of the Moon?
	Analysis	13	Does the moon always stay in the same place?
5th	Knowledge	12	How does a solar eclipse happen?
	Understand	14	Why does the shape of the Moon keep changing?
	Application	12	How does our rotation around the Sun affect the seasons of the year?
	Analysis	14	Why does the moon sometimes appear big and sometimes small?
6th	Knowledge	15	Do the other planets in the solar system revolve around themselves like Earth?
	Understand	15	Why is the Moon in different phases?
	Application	12	Why does the moon sometimes appear during the day as well?
	Analysis	12	Do other planets in the solar system also have moons like the Moon?

## Appendix H

**Table 8: Thematic Findings of Control Group Students' Feedbacks**

Theme	Category	Code	Explanation	
Visualization and Modeling Challenges	Visual Comprehension Difficulties	Geometric shape and rotational motion of the Sun	Difficulty imagining and understanding the shape and motion of the Sun	
		Surface structure and atmosphere of the Moon	The complexity of understanding the Moon's surface and atmosphere, the difficulty of expressing thoughts on the possibility of life	
		Models of movement of celestial bodies	Uncertainty in modeling the relative motions between the Sun, Earth and Moon	
		Modeling and Implementation Challenges	Size comparison of the Sun and Earth	Inability to model the difference in magnitude between the Sun and the Earth
		Rotation of the Moon and its phases	Incomprehension of the relationship between the Moon's rotation around the Earth and its phases	
Lack of Interaction and Support	Individual Understanding and Doubts	Lack of comprehension security	Insecurity about fully comprehending the material being read	
		The need for interactive support	The need to ask questions and get feedback	