




## The effect of e-argumentation supported instruction on middle school students' academic achievement, argumentation skills and epistemic beliefs


Eralp Bahcivan, Nuh Yavuzalp, Sevgi Kingır, Mehmet Demirbağ, Erhan Güneş, Mutlu Tahsin Üstündağ, Muhsin Kılıç, Yunus Özdemir & Niyazi Öztürk

To cite this article: Eralp Bahcivan, Nuh Yavuzalp, Sevgi Kingır, Mehmet Demirbağ, Erhan Güneş, Mutlu Tahsin Üstündağ, Muhsin Kılıç, Yunus Özdemir & Niyazi Öztürk (14 Oct 2025): The effect of e-argumentation supported instruction on middle school students' academic achievement, argumentation skills and epistemic beliefs, International Journal of Science Education, DOI: [10.1080/09500693.2025.2574527](https://doi.org/10.1080/09500693.2025.2574527)

To link to this article: <https://doi.org/10.1080/09500693.2025.2574527>


 Published online: 14 Oct 2025.

 Submit your article to this journal [↗](#)

 Article views: 315






 View related articles [↗](#)

 View Crossmark data [↗](#)

 Citing articles: 1 View citing articles [↗](#)



# The effect of e-argumentation supported instruction on middle school students' academic achievement, argumentation skills and epistemic beliefs

Eralp Bahcivan <sup>a</sup>, Nuh Yavuzalp <sup>a</sup>, Sevgi Kingir<sup>b</sup>, Mehmet Demirbağ <sup>c</sup>, Erhan Güneş <sup>d</sup>, Mutlu Tahsin Üstündağ <sup>e</sup>, Muhsin Kılıç<sup>a</sup>, Yunus Özdemir<sup>a</sup> and Niyazi Öztürk <sup>e</sup>

<sup>a</sup>Faculty of Education, Bolu Abant İzzet Baysal University, Bolu, Türkiye; <sup>b</sup>Faculty of Education, Hacettepe University, Ankara, Türkiye; <sup>c</sup>Faculty of Education, Bursa Uludağ University, Bursa, Türkiye; <sup>d</sup>Faculty of Education, Kırşehir Ahi Evran University, Kırşehir, Türkiye; <sup>e</sup>Faculty of Education, Gazi University, Ankara, Türkiye

## ABSTRACT

This study investigated e-Argumentation software's impact on middle school students' (N=838, grades 5-8) academic achievement, argumentation skills, and epistemological beliefs, compared to Software-free Argumentation and traditional teaching. A mixed-methods design, incorporating a quasi-experimental pretest-posttest control group approach and a multiple case study, assigned students to e-Argumentation, Software-free Argumentation, or Control groups. Data were collected via achievement tests, argumentation rubrics, an epistemological belief scale, an epistemic vignette, and student interviews.

Quantitative analyses revealed that both argumentation groups significantly outperformed controls in academic achievement. The e-Argumentation group demonstrated higher achievement than the no-software group in grades 6 and 8, while the reverse was observed in grade 5; no significant difference was found in grade 7. Descriptively, e-Argumentation groups produced higher quality arguments. The impact on epistemological beliefs was limited; quantitative measures indicated a small e-Argumentation advantage over controls in grades 7-8, though interviews suggested increased student awareness of their beliefs and the nature of scientific knowledge.

In conclusion, argumentation-based instruction effectively improves academic achievement and argumentation skills in science. e-Argumentation software shows promise as a supportive tool, but fostering changes in epistemological beliefs appears to be a more complex, long-term process.

## ARTICLE HISTORY

Received 25 June 2025  
Accepted 17 September 2025

## KEYWORDS

Argumentation; e-argumentation; science education; academic achievement; argumentation skills; epistemological beliefs; middle school students

## Introduction

In science education, argumentation plays a pivotal role, not only enabling students to develop a profound understanding of scientific concepts but also fostering essential

twenty-first-century skills, including scientific process skills, critical thinking, and science literacy (Driver et al., 2000; Erduran & Jimenez-Aleixandre, 2007; Osborne et al., 2004). Argumentation, as a process wherein scientific knowledge is socially constructed, claims are substantiated by evidence and justification, and opposing viewpoints are critically evaluated (Bricker & Bell, 2008; van Eemeren & Grootendorst, 2004), encourages students' active engagement in scientific practices and forms the bedrock of scientific discourse (Cavagnetto, 2010; Chin & Osborne, 2010). Numerous studies have demonstrated that argumentation-based learning environments enhance students' conceptual understanding (Asterhan & Schwarz, 2007; Newton et al., 1999), prove effective in rectifying misconceptions (Asterhan & Schwarz, 2009), and support the development of higher-order thinking skills (Kuhn & Udell, 2003). In this context, argumentation holds the potential to both elevate students' academic achievement and refine their scientific thinking capabilities.

However, a prevalent finding in the literature is that students encounter significant challenges in constructing scientific arguments and actively participating in this process. Students often struggle to support their claims with appropriate and sufficient evidence, establish logical connections (reasoning) between evidence and claims, and critically assess counterarguments (Krajcik et al., 1998; McNeill & Krajcik, 2008; Sandoval & Millwood, 2005). These difficulties can curtail students' capacity to comprehend scientific knowledge and engage in scientific processes, thereby adversely affecting their science learning (Kuhn, 2005).

To address these challenges and enhance the efficacy of the argumentation process, the utilisation of educational technologies is gaining increasing prominence (Fischer et al., 2013; Tsai et al., 2012). Technology-supported argumentation environments can furnish students with various scaffolds for structuring their arguments, organising evidence, interacting with peers, and receiving feedback (Belland, 2010; Sandoval & Reiser, 2004). The international literature features general-purpose argumentation software such as Belvedere, Argumentative, Digalo, Argonaut, and Rationale, alongside various software tailored for science teaching, including ExplanationConstructor (Sandoval & Millwood, 2005; Sandoval & Reiser, 2004), CyberTracker (Ansell & Koenig, 2011), Flyer (Laru et al., 2012), and Zydeco (Delen, 2014). Nevertheless, many of these software exhibit limitations, such as restricted teacher intervention, a common focus on a singular argumentation model, insufficient support for group work, or an inability to fully cater to the specific needs of diverse educational contexts. For instance, a considerable portion of existing software may not provide comprehensive and integrated solutions for students' key difficulties, such as formulating evidence-based explanations, transforming data into claims, or developing sound reasoning.

Beyond the general limitations of current software and students' common difficulties in argumentation, the effectiveness and applicability of an educational tool are intrinsically linked to the consideration of the unique needs of specific educational settings and users during its development. The e-Argumentation software, which underpins this study, is a web-based tool developed with this understanding. It considers the aforementioned limitations and student difficulties documented in the literature, as well as the middle school science curriculum and student needs in Türkiye. The e-Argumentation software employed in this research is distinguished by its flexible structure, capable of supporting diverse argumentation approaches categorised by Cavagnetto (2010) (e.g.

Analytical model, Argumentation-Based Science Learning [ABSL]/Science Writing Heuristic [SWH], and Socioscientific Issues [SSI]). It also permits teachers to intervene with instant feedback and allows students to enrich their arguments with multimedia tools (text, graphics, images, etc.). This software aims to differentiate itself from existing tools by providing structured support to overcome the primary challenges students encounter in the argumentation process (e.g. using evidence, reasoning, and questioning the reliability of evidence sources). Therefore, this study will contribute to the literature by developing a unique e-Argumentation software that overcomes specific pedagogical and technical limitations identified in existing international software (e.g. a focus on a single argumentation model, the lack of real-time teacher intervention), and by examining the holistic effects of this software on students' academic achievement, argumentation skills, and epistemological beliefs, thereby filling the gap in technology-supported argumentation research within the Turkish context.

The theoretical framework of this study is grounded in the central role of argumentation in students' learning processes. Argumentation, by its nature, cognitively activates students, compelling them to support and justify their claims with evidence. This process encourages students to question their existing knowledge, forge connections with new information, and consequently, enhance their academic achievement (Driver et al., 2000; Osborne et al., 2004). Concurrently, participation in the argumentation process directly enables students to cultivate fundamental argumentation skills such as claim construction, evidence presentation, justification, and refutation (Erduran et al., 2004). Furthermore, individuals' epistemological beliefs – their convictions regarding the nature of knowledge and knowing – are considered to be intricately related to their argumentation skills and their propensity to engage in the argumentation process (Duschl & Osborne, 2002; Kuhn, 1993; Nussbaum & Bendixen, 2003; Sandoval & Millwood, 2007). Argumentation can foster the development of more sophisticated epistemological beliefs by creating an environment where epistemic dimensions – such as the nature of knowledge (its certainty and immutability), its source, and its justification – are interrogated (Demirbağ & Bahçivan, 2022; Sampson & Clark, 2006). Therefore, this study holistically examines these three principal student outcomes (academic achievement, argumentation skills, and epistemological beliefs) to investigate the comprehensive impact of an e-Argumentation-supported learning environment.

While the benefits of technology-supported argumentation are increasingly recognised internationally (Fischer et al., 2013), a significant research gap exists within the Turkish educational context. Previous national studies on argumentation have primarily focused on pedagogical development without integrating a technological software component to support teacher and student interaction. The few existing software applications are very limited and generally not designed for direct classroom use or lack crucial features like teacher intervention. This study addresses this gap by developing and systematically evaluating an e-argumentation tool specifically tailored for the needs of Turkish middle school students and the national curriculum. Furthermore, although prior research has often investigated the effects of argumentation on academic achievement, argumentation skills, or epistemological beliefs in isolation, this study offers a novel contribution by holistically examining the interconnected effects on all three variables simultaneously within a single technology-supported environment. Given the theoretical links suggesting that the argumentation process enhances conceptual understanding,

directly fosters argumentation skills, and is deeply intertwined with epistemic development, a comprehensive investigation of these outcomes provides a more robust understanding of the overall impact of such an intervention. Accordingly, the objective of this article is to examine the effects of argumentation instruction supported by e-Argumentation software on the academic achievement, argumentation skills, and epistemological beliefs of middle school students (grades 5, 6, 7, and 8), comparing it with argumentation instruction without software support and traditional science teaching methods.

## Theoretical framework

This study is predicated on the central role of argumentation in science education and its interactive effects on students' academic achievement, argumentation skills, and epistemological beliefs. These core concepts and the theoretical relationships among them are elaborated below.

### *Argumentation*

Argumentation, in its broadest sense, is a social and cognitive process wherein claims are supported, questioned, and evaluated using evidence and reasoning (Van Eemeren & Grootendorst, 2004). Within the context of science education, argumentation is regarded as a fundamental practice through which scientific knowledge is generated, appraised, and communicated (Driver et al., 2000; Duschl & Osborne, 2002). Simon et al. (2006) define argumentation as a process where the essential components of an argument – namely claim, data, warrant, and backing – are coherently interconnected. McNeill and Krajcik (2007), in contrast, conceptualise it as a process that integrates the components of claim, evidence, and reasoning. Scientists endeavour to make sense of the natural world by posing questions, gathering evidence pertinent to these questions, formulating claims through the interpretation of evidence, and contributing to the advancement of knowledge by debating these claims within the scientific community (Bricker & Bell, 2008). Consequently, argumentation is not merely a mode of communication but an epistemic practice central to scientific thinking and knowledge construction (Cavagnetto, 2010). The incorporation of argumentation in science classrooms enables students to achieve a more profound understanding of scientific concepts, develop scientific process skills, and attain a better grasp of the nature of science (Erduran & Jimenez-Aleixandre, 2007). This study employed three distinct argumentation approaches to ensure a comprehensive implementation (Cavagnetto, 2010). The Analytic Model involves the explicit teaching of argument components (e.g. claims, data). Inquiry-based models (e.g. ABSL) integrate argumentation implicitly within the scientific inquiry process, where students generate claims from their own data. The Socioscientific Issues (SSI) approach uses controversial, dilemma-based topics to foster reasoning skills. The rationale for employing these diverse models was to move beyond a single-method limitation often seen in practice. This methodological flexibility aimed to enrich the implementation of argumentation and was a core design feature of the e-Argumentation software, which supports all three approaches.

### **Argumentation skills**

Argumentation skills encompass a range of competencies, including the ability to formulate a claim, substantiate that claim with valid and reliable evidence, establish logical warrants between evidence and claim, critically evaluate opposing claims and evidence, and effectively defend one's own arguments. Toulmin's (1958) model of argument offers a frequently utilised framework for analyzing these skills, delineating the core components of an argument as claim, data, warrant, backing, qualifier, and rebuttal. Students' proficiency in understanding and effectively employing these components serves as a significant indicator of their scientific literacy (Osborne et al., 2004). However, the literature consistently indicates that students encounter difficulties, particularly in utilising evidence, constructing robust warrants between claims and evidence, and considering counterarguments (Krajcik et al., 1998; McNeill & Krajcik, 2008; Sandoval & Millwood, 2005). The cultivation of these skills is of paramount importance, not only for success in science classes but also for enabling students to make informed decisions regarding the complex situations they encounter in their daily lives (Kuhn, 1991).

### **Epistemological belief and reasoning**

Epistemological beliefs refer to individuals' personal theories and assumptions concerning the nature of knowledge and the processes of knowing, encompassing aspects such as what knowledge is, how it is acquired and justified, and its degree of certainty (Hofer & Pintrich, 1997; Schommer, 1990). These beliefs can influence individuals' learning approaches, problem-solving strategies, and their engagement in argumentation processes (Nussbaum & Bendixen, 2003). Kuhn (1991, 1993) categorised levels of epistemological development as absolutist, multiplist, and evaluativist. Individuals at the absolutist level perceive knowledge as an absolute and authoritatively derived structure, whereas those at the multiplist level consider all views to be equally valid. In contrast, individuals at the evaluativist level acknowledge that knowledge can be assessed based on evidence and arguments, and that some perspectives may hold greater validity than others.

A bidirectional relationship is posited between argumentation and epistemological beliefs. On one hand, individuals' epistemological beliefs influence how they construct arguments, evaluate evidence, and their openness to conflicting viewpoints (Duschl & Osborne, 2002; Sandoval & Millwood, 2007). For instance, students possessing more sophisticated (evaluativist) epistemological beliefs are reportedly capable of producing more complex and higher-quality arguments (Mason & Scirica, 2006). On the other hand, argumentation-based learning environments can enhance students' epistemological awareness and reasoning skills by exposing them to diverse perspectives and encouraging them to question evidence and justify their own thinking (Demirbağ & Bahçivan, 2022; Sampson & Clark, 2006). Epistemic reasoning, in this context, denotes the process by which individuals evaluate knowledge claims concerning a specific topic or problem, compare disparate sources of evidence, and arrive at reasoned judgments predicated on these evaluations.

### **Academic achievement**

Academic achievement, within the scope of this study, pertains to students' level of understanding of fundamental concepts related to a specific topic or unit in science

courses, their ability to apply these concepts in problem-solving, and their capacity to interpret scientific principles. Argumentation-based instruction is believed to positively influence students' academic achievement by encouraging them to forge in-depth connections between concepts, actively process information, and construct their own understanding, rather than merely memorising superficial information (Asterhan & Schwarz, 2007; Zohar & Nemet, 2002). When constructing and discussing arguments, students are afforded opportunities to utilise scientific concepts in varied contexts, clarify their own thinking, and identify and rectify potential misconceptions. These cognitive processes can contribute to more durable and meaningful learning, thereby enhancing academic achievement.

This study investigates the intricate relationships among argumentation skills, epistemological beliefs, and academic achievement within the framework of a technology-supported argumentation environment. While the argumentation process itself provides a foundation for students to develop both their argument construction skills and their academic achievement, the epistemological stance and reasoning skills they exhibit during this process can also influence their learning outcomes. A primary focus of this study is to determine how technology-supported environments shape and facilitate these interactions.

Moreover, this three-group research design was intentionally chosen to disentangle the effects of the argumentation pedagogy itself from the added value of the technology-supported environment, using existing (traditional) instruction as a baseline for comparison. In line with the aforementioned purpose, this study sought to answer the following research questions:

1. How does the use of e-Argumentation software in science lessons affect the development of middle school students' argumentation skills (argument quality) compared to argumentation practices without software support and traditional teaching?
2. How does the use of e-Argumentation software in science courses affect middle school students' academic achievement levels in science compared to argumentation practices without software support and traditional teaching?
3. How does the use of e-Argumentation software in science courses affect middle school students' epistemological beliefs compared to argumentation practices without software support and traditional teaching?

## **Method**

### ***Research design***

This study employed a mixed-methods research design to comprehensively examine the effects of e-Argumentation-supported instruction, argumentation instruction without software, and traditional science instruction on middle school students' academic achievement, argumentation skills, and epistemological beliefs. The quantitative dimension of the study utilised a quasi-experimental design with a pretest-posttest control group, facilitating comparisons between groups. For the qualitative dimension, a multiple case study approach was adopted to deepen and enrich the quantitative findings

derived from the quasi-experimental study and to gain a more nuanced understanding of the participants' experiences. This approach permitted inter-group comparisons using quantitative data while concurrently employing qualitative data to explore changes in students' argumentation skills and epistemic reasoning in greater depth (Creswell, 2007; Creswell & Plano Clark, 2011).

### **Participants**

This study was conducted with science teachers and the students in their classrooms who volunteered to participate in a more extensive research project. The participants comprised 25 science teachers from 13 different middle schools in Bolu, Türkiye, and their students. A total of 838 middle school students (360 from 5th grade, 158 from 6th grade, 183 from 7th grade, and 137 from 8th grade) were included in the study. The assignment of the 25 participating teachers into the two experimental conditions (e-Argumentation,  $n = 12$ ; Software-free Argumentation,  $n = 13$ ) was conducted through a structured matching procedure during the professional development phase. This procedure was based on quantitative data obtained from in-class observation forms, such as the 'Assessment of Scientific Argumentation in the Classroom (ASAC)' observation protocol, which were completed during the trial implementations. The primary goal of this matching was to ensure that the two experimental groups were comparable in terms of their baseline argumentation pedagogy skills before the main intervention began. Within this matching framework, teachers' preferences for using the e-Argumentation software were also accommodated, particularly for the five teachers who participated with classes from two different grade levels, four of whom opted to use the software in both. Subsequently, from each teacher's two classes with similar characteristics at the same grade level, one intact class was randomly assigned to the experimental condition corresponding to the teacher's group (e-Argumentation or Software-free Argumentation), and the other was assigned to the control group. Specifically, statistical analysis of the pre-test scores from the Conceptual Achievement Test and the Epistemological Belief Scale confirmed that there were no statistically significant pre-existing differences between the two classes taught by each teacher before one was randomly assigned to the experimental condition and the other to the control group. The randomisation, therefore, occurred at the class level, not at the individual student level. Consequently, students were placed into one of three groups based on their teacher's assigned teaching method and their class assignment:

- **e-Argumentation Group:** The students of the 12 teachers in this group engaged in argumentation-based activities using the web-based e-Argumentation software developed for this study.
- **Software-free Argumentation Group:** The students of the 13 teachers in this group conducted argumentation-based activities without software support, utilising methods such as worksheets and in-class discussions.
- **Control Group:** This group consisted of students from the other classrooms of the teachers in both experimental groups where no intervention was applied. In these classes, teachers continued to implement their existing (traditional) curriculum and methodologies.

To control for the teacher variable, each teacher taught both an experimental group (e-Argumentation or Software-free Argumentation) and a control group at the same grade level. The distribution of participant students by group and grade level is detailed in Table 1.

For the qualitative dimension of the study, a total of 80 students participated from the experimental groups. The sampling strategy was purposive in that it sought to gather insights from every unique experimental classroom context. The specific selection method was volunteer sampling. From each experimental class (both e-Argumentation and Software-free Argumentation), at least two students who expressed willingness to participate were included in the epistemic vignette application and semi-structured interviews. This approach ensured that the qualitative data reflected experiences across the full range of implementation settings.

### **Implementation process and e-argumentation software**

Prior to implementation, teachers in the experimental groups (e-Argumentation and Software-free Argumentation) participated in professional development activities (Professional Learning Community – PLC). These activities focused on the role of argumentation in science education, its theoretical underpinnings, various argumentation models (e.g. Analytic Model, Argumentation-Based Science Learning [ABSL]/Science Writing Heuristic [SWH], Socioscientific Issues [SSI]), and classroom application strategies. These PLCs comprised collective interviews and training meetings conducted over nine sessions presented in Table 2.

According to Table 2, each PLC session took at least one week, 8 months in total. Teachers in the e-Argumentation group also received supplementary training on the utilisation of the e-Argumentation software. All teachers were provided with on-the-job support throughout the implementation period (two months between PLC6 and

**Table 1.** Distribution of Participant Students by Grade Level and Groups.

Class Level	Group Type	Number of Branches	Number of Students
<b>Grade 5</b>	e-Argumentation (Experiment)	7	91
	e-Argumentation (Control)	7	90
	Software-free Argumentation (Experiment)	6	94
	Software-free Argumentation (Control)	6	85
	<b>Total Grade 5</b>	<b>26</b>	<b>360</b>
<b>Grade 6</b>	e-Argumentation (Experiment)	3	39
	e-Argumentation (Control)	3	37
	Software-free Argumentation (Experiment)	3	39
	Software-free Argumentation (Control)	3	43
	<b>Total Grade 6</b>	<b>12</b>	<b>158</b>
<b>Grade 7</b>	e-Argumentation (Experiment)	4	60
	e-Argumentation (Control)	4	64
	Software-free Argumentation (Experiment)	2	29
	Software-free Argumentation (Control)	2	30
	<b>Total Grade 7</b>	<b>12</b>	<b>183</b>
<b>Grade 8</b>	e-Argumentation (Experiment)	2	37
	e-Argumentation (Control)	2	39
	Software-free Argumentation (Experiment)	3	37
	Software-free Argumentation (Control)	3	24
	<b>Total Grade 8</b>	<b>10</b>	<b>137</b>
<b>General Total</b>		<b>60</b>	<b>838</b>

PLC7, and four months between PLC8 and PLC9, with support provided at least monthly if required). Furthermore, to ensure fidelity to the intended pedagogical approaches in each condition, the research team conducted regular classroom observations across all groups. In the Software-free Argumentation classes, researchers used a pre-prepared checklist to verify that the core components of the argumentation framework were being implemented and provided on-the-job support to teachers as needed. For the control group, observations and field notes served to monitor and ensure fidelity by confirming that teachers adhered to their traditional methods and did not introduce argumentation practices learned during the PLCs, thus preventing treatment contamination. The assignment of teachers to the experimental groups was directly linked to their participation in this PLC process. To mitigate potential selection bias, a matching procedure was employed after the initial PLC sessions (specifically at PLC8). Data from in-class observation forms, collected during the trial implementations that were part of the initial PLC training, were used to form two experimental groups with comparable baseline argumentation pedagogy skills. This process ensured that the e-Argumentation and Software-free Argumentation groups were pedagogically similar before the main intervention began.

### *Features of e-argumentation software*

The e-Argumentation software used in this study was developed under TUBITAK 1001 project number 219K028. It is a web-based platform designed to structure, facilitate, and support argumentation processes in science courses. The development process considered several factors: student difficulties frequently highlighted in the literature (e.g. using evidence, reasoning; Krajcik et al., 1998; McNeill & Krajcik, 2008; Sandoval & Millwood, 2005); teachers' expressed needs for tools to simplify argumentation practices and enhance student engagement; known limitations of existing argumentation software (e.g. restricted teacher intervention, focus on a single model, lack of interaction); and the specific characteristics of the educational environment in Türkiye. The main features of the software are as follows:

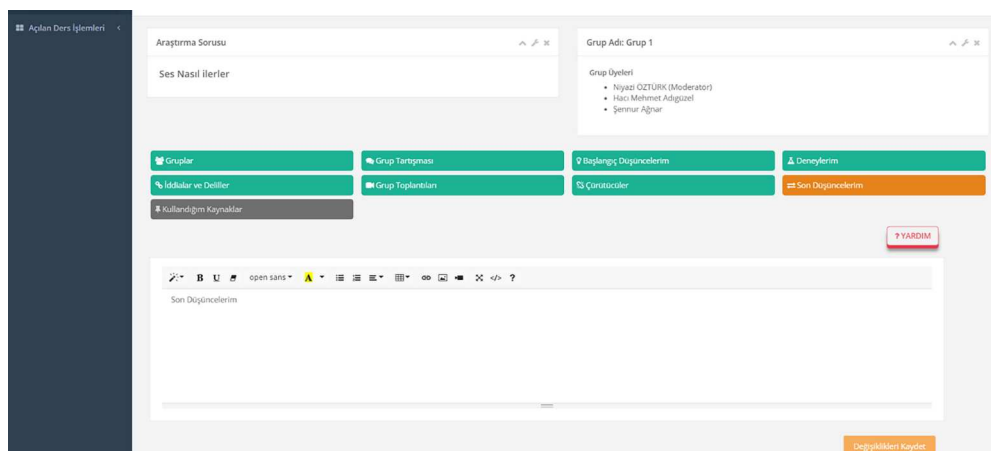
- **Flexible Argumentation Models:** The software is designed with a modular structure that can adapt to various argumentation approaches commonly employed in science education (e.g. the Analytic Model; inquiry-based models such as

**Table 2.** Overview of Teacher Professional Learning Community (PLC) Sessions.

Session	Main Purpose/Topic	Duration
PLC 1	Introduction to project goals and team members.	1 Week
PLC 2	Identifying participants' current teaching and learning conceptions.	1 Week
PLC 3	Defining argumentation and introducing different argumentation approaches.	1 Week
PLC 4	Preparing and implementing argumentation-oriented lesson plans with team support.	2 Weeks
PLC 5	Evaluating the initial classroom implementations.	1 Weeks
PLC 6	Developing solutions for problems and sharing examples of good practice.	2 Months
PLC 7	General reflections on the trial implementation period.	1 Week
PLC 8	Official start of the experimental study; separation of teachers into e-Argumentation and Software-free Argumentation groups and differentiation of training support.	4 Months
PLC 9	Collecting final reflections on the entire process.	1 Week

Argumentation-Based Science Learning [ABSL]/Science Writing Heuristic [SWH]; and discussions within the context of Socioscientific Issues [SSI]). Teachers can select the most appropriate argumentation structure based on the lesson's objectives and topic.

- **Structured Activity Flow:** Students typically begin by encountering an initial problem or question, after which they articulate their individual prior knowledge or predictions. They then collaborate with their group to formulate a claim. The software provides a dedicated section for adding evidence to support this claim. In this section, students can input text, upload files such as images or videos from their computers or the internet, or utilise data visualisation tools like tables and graphs. They are also guided to write their reasoning, thereby linking the evidence to their claim. (See [Figure 1](#) for the claim and evidence submission screen).
- **Interaction and Discussion:** In addition to constructing their own arguments, students can view arguments uploaded by other groups to the system. The software enables groups to comment on each other's claims, evidence, or reasoning, ask questions, and present counterclaims or rebuttals. These interactions can occur synchronously or asynchronously. (See [Figure 2](#) for the group discussion interface).
- **Teacher Role and Intervention:** The software empowers the teacher to actively monitor and intervene in the process, rather than merely initiating and concluding it. Teachers can instantly view the work of all groups, provide written feedback directly through the software when students encounter difficulties (e.g. insufficient evidence, flawed reasoning), and offer additional questions or prompts. This feature aims to deliver the instant teacher support that is noted as lacking in some argumentation software discussed in the literature (Akpınar et al., 2015). (See [Figure 3](#) for the student feedback interface).
- **Supports for Student Difficulties:** The software incorporates features designed to alleviate common difficulties students experience during the argumentation process.



**Figure 1.** Screenshot of the claim and evidence submission interface.

The screenshot displays a web-based interface for group discussions. At the top, there are navigation buttons: "Görüntüleme Seçenekleri" (View Options) with sub-buttons for "İddialar" (Claims), "Yenile" (Refresh), "Değiller" (Not), "Karşı İddialar" (Counter-claims), and "Yerleşim Düzeni Seçenekleri" (Layout Options). The main content area is divided into six panels, each representing a group (Grup 1 to Grup 6). Each panel contains a task description, a table of data, and a "Kargıt İddia" (Counter-claim) button. The tasks involve discussing the factors affecting the rate of evaporation of a liquid, with data tables showing the volume of liquid and the time taken for it to evaporate under different conditions.

**Grup 1:** Task: "Isınma miktarına etki eden faktörlerin maddenin miktarı, verilen ısı ve öz ısısı olduğunu düşünüyorsunuz." Data table: 

Sıvının Cinsi	0 dk	1 dk	2 dk	3 dk
Zeytin Yağı	24	29	32	37
Su	23	26	29	33

**Grup 2:** Task: "Maddenin yoğunluğuyla maddenin ısınma miktarının alakası yoktur. Önemli olan maddenin öz ısısıdır." Data table: 

sıvı cinsi	0dk	1dk	2dk	3dk
Su	24	26	29	34

**Grup 3:** Task: "maddenin cinsi ısınma miktarına etki eder." Data table: 

	0 dk	1 dk	3 dk	5 dk
Zeytinyağı	24	25	30	36
Su	22	23	28	33

**Grup 4:** Task: "Yağın öz ısısı suyun öz ısısından daha küçük olduğu için daha fazla ısınmıştır. Az olduğu için değil." Data table: 

Sıvının Cinsi	0 dk	1 dk	2 dk	3 dk
Zeytin Yağı	24	29	32	37
Su	23	26	29	33

**Grup 5:** Task: "çünkü yaptığımız deneyde gördük ki zeytinyağı daha fazla ısı almış bu da demek oluyor ki sıvının yoğunluğu ısı miktarını etkiler" Data table: 

Su	24	26	29	34
Su	24	26	29	34

**Grup 6:** Task: "Yağ ve suyu eşit ısıtıcılarda eşit süre ısıttığımızda öz ısısı büyük olan su daha geç ısındı. Öz ısısı küçük olan yağ suya göre daha erken ısınarak suya göre sıcaklığını daha fazla artırdı. Öz ısısı az olan daha erken ısınır daha erken soğur. Öz ısısı daha büyük olan geç ısınır daha geç soğur..."

Figure 2. Screenshot of the group discussion interface.

For example, it encourages students to seek and utilise diverse types of evidence beyond personal experiences, such as experimental results, research data, and expert opinions. It also assists them in organising and interpreting data in various formats, such as tables and graphs, and provides indirect prompts to question the reliability of evidence.

- **Accessibility and Usability:** Being web-based and compatible with various devices (smart boards, computers, tablets), the software is suitable for use in diverse learning environments (face-to-face, distance, hybrid). Usability principles were a key consideration in the interface design.

The screenshot shows a "Geri Bildirim Ver" (Give Feedback) form. The form has a table with columns for "İçerik Tipi" (Content Type), "İçerik" (Content), and "Tarih" (Date). Below the table is a text input field for "Geri Bildiriminiz:" (Your Feedback) and a "Kaydet" (Save) button.

İçerik Tipi	İçerik	Tarih
İlk Yazılan	fosit yakıtlardan fazlaca kullanımı sonucu oluşur.Asıt yağmurları asitlik kimyasalların hava olayları yoluyla yerlere ulaşmazdır.Çevreye fazlaca zarar verir ve biz asıt yağmurlarını gözle göremeyiz.Yenilebilir enerji kaynakları kullanmak asıt yağmurlarının oluşumunu azaltır.	2022-01-13 16:01:49:530
Verdiğiniz Geri Bildirim	hava olayları ile nasıl yağmurlara karşılıyor.. Açıklayabilir misiniz.	2022-01-13 16:10:26:400
Geri Bildirim Sonucu Düzeltilen	fosit yakıtlardan fazlaca kullanımı sonucu oluşur.Asıt yağmurları asitlik kimyasalların hava olayları yoluyla yerlere ulaşmazdır.Çevreye fazlaca zarar verir ve biz asıt yağmurlarını gözle göremeyiz.Yenilebilir enerji kaynakları kullanmak asıt yağmurlarının oluşumunu azaltır.	2022-01-13 16:11:47:127
Verdiğiniz Geri Bildirim	fosit yakıtlar nelerdir açıklayabilir misiniz.?	2022-01-13 16:17:52:940
Geri Bildirim Sonucu Düzeltilen	fosit yakıtlardan fazlaca kullanımı sonucu oluşur.Asıt yağmurları asitlik kimyasalların hava olayları yoluyla yerlere ulaşmazdır.Çevreye fazlaca zarar verir ve biz asıt yağmurlarını gözle göremeyiz.Yenilebilir enerji kaynakları kullanmak asıt yağmurlarının oluşumunu azaltır.	2022-01-13 16:18:08:860

Figure 3. Screenshot of the student feedback interface.

### **Experimental implementation**

For each of the 5th, 6th, 7th, and 8th grade levels included in the study, instructional processes were conducted within the scope of different science units. These units were specifically identified by the teachers for their respective grade levels during the Professional Learning Community (PLC) process. The selected units were taught concurrently in all experimental (e-Argumentation, Software-free Argumentation) and control groups at each grade level, targeting the same fundamental learning outcomes. The instructional time for each unit was planned in accordance with the course hours allocated in the curriculum for that unit, and the implementations continued for one academic semester. Prior to the experimental application, all students in all groups were administered the relevant measurement tools as a pre-test. The same tools were administered as a post-test at the conclusion of the unit. The specific science units selected for the intervention were as follows: ‘UNIT-6: Human and Environment’ for the 5th grade, ‘UNIT-5: Sound and Its Properties’ for the 6th grade, ‘UNIT-6: Reproduction, Growth, and Development in Living Things’ for the 7th grade, and ‘UNIT-4: Matter and Industry’ for the 8th grade. The general framework of the teaching processes carried out in each group is presented below:

- **e-Argumentation Group Implementation:** Lessons in this group were conducted under teacher guidance using the e-Argumentation software detailed above. Students progressed through the software’s modular structure, encountering a research question or problem scenario. They expressed their initial individual thoughts, collaborated with their groups to formulate claims, uploaded evidence in textual, visual, or data formats via the software to support these claims, wrote their justifications, examined their peers’ arguments, engaged in online discussions, and received feedback from their teachers. To ensure treatment fidelity, this core instructional sequence involving the software’s primary modules was consistently applied across all classes in the e-Argumentation group. The standardised process for all participating students included: (1) encountering a problem scenario, (2) recording initial individual thoughts, (3) collaboratively formulating a group claim, (4) submitting supporting evidence and reasoning, and (5) engaging in inter-group online discussions. Likewise, the ability for teachers to actively monitor student progress and provide real-time feedback through the software was an integral and consistently utilised component of the pedagogy in all e-Argumentation classrooms.
- **Software-free Argumentation Group Implementation:** Lessons in this group were conducted within the framework of argumentation approaches (e.g. ABSL, SWH) that teachers had learned in the PLCs, but without the use of software. In this process, students typically began with a research question or scenario. This was followed by data collection (conducting experiments, making observations, or examining provided data sets), data analysis, claim formulation, and the substantiation and justification of these claims with evidence. They presented and discussed their constructed arguments using structured worksheets, poster presentations, or through in-class small-group and large-group discussions.
- **Control Group Implementation:** In this group, teachers taught the relevant units using their accustomed traditional methods, which were generally teacher-centered and textbook-based. These methods included lectures, question-answer sessions

(primarily focused on recall), and the solving of activities and problems from the textbook. Argumentation processes, wherein students formulate their own claims, support them with evidence, or engage in in-depth discussions with peers, were not incorporated.

To minimise implementation variability and ensure fidelity across the experimental conditions, several standardisation measures were taken. Although teachers were not provided with a rigid, verbatim script, all experimental group teachers: (1) received the same comprehensive PLC training on argumentation pedagogy, (2) collaboratively developed their argumentation-oriented lesson plans with the support of the research team during the PLCs, and (3) taught the same pre-determined science units targeting the same fundamental learning outcomes for their respective grade levels. The instructional process for the e-Argumentation group was further standardised by the software's inherent structured workflow, while the Software-free Argumentation group used common materials such as structured worksheets to guide their activities.

### **Data collection tools**

The data collection tools utilised to measure students' academic achievement, argumentation skills, and epistemological beliefs are detailed in [Table 3](#).

- **Conceptual Achievement Test:** For each grade level (5-8), achievement tests were developed for this study. These tests consisted of multiple-choice and open-ended questions derived or adapted from previous national examinations (e.g. National Central Exams), international assessments (e.g. PISA, TIMSS), and national achievement tests, all aligned with the learning outcomes of the relevant units. To ensure content validity, the created question pool was reviewed by three subject matter experts (science educators) within the project team. The validation process involved matching each test item to the targeted learning outcomes using a table of specifications. Following the expert review, it was concluded that content validity was established and no new items were needed. After a pilot study, Cronbach's alpha reliability coefficients were calculated as .70 for 5th grade, .72 for 7th grade, and .83 for 8th grade. The coefficient for the 6th-grade test was .69, which is considered acceptable for the purposes of this study. The slightly low reliability of 6th-grade test scores may be attributed to the nature of the content being measured.
- **Argumentation Skills Assessment Rubrics:** Students' argumentation skills were evaluated based on the arguments they produced. In the e-Argumentation group, this involved assessing the arguments (claims, evidence, justifications, etc.) that groups entered into the e-Argumentation software either during or at the end of the application. In the Software-free Argumentation group, skills were evaluated through written arguments created using materials such as structured worksheets and poster presentations. The rubrics used were appropriate to the specific argumentation approach applied. For ABSL and SWH approaches, a rubric adapted by Demirbağ and Günel (2014) from Choi's (2008) study was utilised. For the Analytic Model and SSI approaches, qualitative categories adapted by Hiğde and Aktamiş (2017) from Osborne et al.'s (2004) argument levels were employed, focusing on the presence

**Table 3.** Student Data Collection Tools Used in the Study and Their Characteristics.

Name of Data Collection Tool	Measured Variable/Objective	Source/ Development	Implementation Time	Evaluation/Analysis Method
Conceptual Achievement Test	Academic Success	It was developed by the project team based on PISA, TIMSS, National Central Exams.	Pre Test - Post Test	Multiple-choice questions were scored with true/false, open-ended questions with rubric. For comparisons between groups, t-test, (M)ANOVA/ (M)ANCOVA.
Argumentation Skills Assessment Rubrics	Argumentation Skills/Argument Quality	For ATBS/ATS: Demirbağ and Günel (2014) based on Choi (2008). For Analytic/SSI: Hiğde and Aktamış (2017) based on Osborne et al. (2004).	Products during/ after application	Qualitative/quantitative scoring according to relevant rubrics/ categories. Descriptive statistics.
Epistemological Belief Scale	Epistemological Beliefs (Nature of Knowledge/ Knowing)	The scale adapted by Conley et al. (2004) / Bahçivan (2014) was shortened and restructured for the project.	Pre Test - Post Test	5-point Likert-type scale scores. (M)ANOVA/ (M)ANCOVA for comparisons between groups.
Epistemic Vignett	Epistemic Reasoning/Profile (Definitivist, Pluralist, Value)	Kuhn et al. (2008) / Bahçivan and Cobern (2016).	Post Test (Case Study Learning)	Qualitative content analysis of open-ended responses, categorisation with Kılıç (2020) rubric.
Student Interviews (Dialogic D. Prot.)	Argumentation experiences, opinions, perception of change	Semi-structured interview protocol developed by the project team.	After Implementation (Case Study)	Qualitative content analysis of interview transcripts, thematic coding.

and quality of components such as claim, evidence, and justification. Due to technical constraints, skill assessments in the e-Argumentation group were conducted at the group level rather than individually.

- **Epistemological Belief Scale:** To measure students' beliefs about the nature of knowledge and knowing (certainty, simplicity, source, justification), a scale originally developed by Conley et al. (2004) and adapted into Turkish by Bahçivan (2014) was utilised. For this study, the scale was condensed to 12 items (Nature of Knowledge: 5 items; Nature of Knowing: 7 items) and presented in a 5-point Likert-type format. This reduction was implemented following expert consultation, with the rationale of preventing student fatigue and burnout, as the scale was administered alongside the conceptual achievement test. The construct validity of this revised form of the scale was supported by confirmatory factor analysis (CMIN/DF = 3.34; CFI = .97; RMSEA = .05), and the reliabilities of its sub-dimensions (Nature of Knowledge  $\alpha = .72$ , Nature of Knowing  $\alpha = .73$ ) were found to be satisfactory. Overall reliability of the test scores was calculated as  $\alpha = .83$ . The scale was administered as a pre-test and a post-test.
- **Epistemic Vignette:** A scenario (vignette), developed by Kuhn et al. (2008) and adapted into Turkish by Bahçivan and Cobern (2016), was used to qualitatively determine students' epistemic reasoning levels and epistemological positions (absolutist, multiplist, evaluativist). The scenario presented two conflicting expert opinions on

the extinction of dinosaurs and asked students to evaluate these views. As part of the validity process for the current study, expert opinion was sought from three science educators, and the instrument was deemed appropriate for use with the target student population. As indicated in the Participants section, this instrument was administered as a post-test to 80 students selected from the experimental groups for the qualitative dimension of the study. Their responses were analyzed using a rubric developed by Kılıç (2020), based on Kuhn et al.'s (2008) levels of epistemological understanding.

- **Student Interviews:** In the qualitative dimension of the study, the same 80 students who participated in the Epistemic Vignette application (as described in the Participants section) were interviewed. These interviews, guided by a semi-structured protocol (Dialogic Exchange Protocol) developed by the project team, explored their argumentation/e-Argumentation experiences, learning processes, and perspectives on scientific knowledge. Given the large number of participants in the qualitative sample ( $N = 80$ ), each interview was intentionally brief, lasting approximately 5 min, to capture students' most immediate reflections following the intervention. With participant consent, these interviews were audio-recorded by project team members and subsequently transcribed. The interview protocol consisted of open-ended questions designed to understand students' general perceptions of the argumentation process, their experiences with the e-Argumentation software (if applicable) or software-free argumentation activities, the challenges they encountered, the contributions of these experiences to their learning, and any potential shifts in their thinking about the nature of scientific knowledge.

### **Data analysis**

SPSS Statistics was used for the analysis of quantitative data (Achievement Test, Epistemological Belief Scale). To ascertain whether significant differences existed between the pre-test scores of the groups and to compare post-test scores, appropriate parametric tests (independent samples t-test, one-way analysis of variance [ANOVA], multivariate analysis of variance [MANOVA], multivariate analysis of covariance [MANCOVA]) were employed. This was done after verifying the assumptions of normality (skewness and kurtosis values) and homogeneity of variances (Levene's test, Box's M test). When significant differences were detected, effect size values (Cohen's  $d$  or eta-square [ $\eta^2$ ]) were calculated to interpret the magnitude of the difference. For Cohen's  $d$ , .20 was interpreted as a small effect, .50 as medium, and .80 as large; for eta-square, .01 was considered small, .06 medium, and .14 large (Green & Salkind, 2014; Pallant, 2005). A direct three-group comparison (e.g. via one-way ANOVA) was not performed due to a key practical constraint of the study: each teacher had only two available class sections per grade level, making it impossible for a single teacher to implement all three conditions. Consequently, each teacher taught one experimental and one control group, serving as their own control. This design necessitated a series of planned pairwise comparisons to rigorously analyze the within-teacher (Experimental vs. Control) and between-teacher (e-Argumentation vs. Software-free) effects, which was deemed the most appropriate analytical strategy.

Content analysis was the chosen method for analyzing qualitative data (Argument Assessments, Vignette Responses, Interviews) (Krippendorff, 2004). Students' written arguments and vignette responses were scored or categorised using relevant rubrics or established categories. The analysis of the qualitative data from the student interviews followed a systematic thematic analysis approach. All interview recordings were first transcribed verbatim. Subsequently, a team of three researchers independently coded the transcripts to identify emerging themes and categories related to students' experiences and perceptions. To ensure reliability, the researchers then met to compare their coding, discuss differences in interpretation, and collaboratively refine the thematic structure until full consensus was reached on all codes and themes. To enhance the validity of the findings, strategies such as data triangulation (utilising different data collection tools) and, in some instances, member checking (participant confirmation) were employed (Creswell, 2007). It is important to note that the argumentation skills data for the e-Argumentation group were analyzed at the group level.

## Findings

This section presents the analysis results concerning the effects of e-Argumentation-supported instruction, argumentation instruction without software support, and traditional instruction on middle school students' academic achievement, argumentation skills, and epistemological beliefs.

### *Findings related to academic achievement*

This sub-section comparatively presents the effects of e-Argumentation-supported instruction, argumentation instruction without software support, and traditional instruction on middle school students' academic achievement in science courses, based on pre-test and post-test scores from the Conceptual Achievement Test. Analyses were conducted separately for each grade level. In all analyses, it was first checked whether there was a significant difference between the pre-test scores of the groups. In cases where no significant pre-test difference was found, post-test scores were compared using an independent samples *t*-test. For significant differences, effect sizes were calculated using Cohen's *d* and interpreted as .20 (small), .50 (medium), and .80 (large) effect (Green & Salkind, 2014).

### *Findings related to the fifth grade level*

- **Comparison of e-Argumentation Group and Control Group:** No significant difference was found between the achievement test pre-test scores of the two groups,  $t(179) = -0.95, p = .342$ . When the post-test scores were compared, students in the e-Argumentation group (Mean = 6.71, SD = 1.74) demonstrated statistically significantly higher achievement than students in the control group (Mean = 6.09, SD = 1.83),  $t(179) = 2.36, p = .019$ . The calculated effect size ( $d = .35$ ) indicated a small difference.
- **Comparison of Software-free Argumentation Group and Control Group:** Similarly, no significant difference was detected between the pre-test achievement scores of these groups,  $t(177) = -0.27, p = .786$ . Analysis of the post-test scores revealed that students

in the Software-free Argumentation group (Mean = 7.67, SD = 1.91) were significantly more successful than students in the control group (Mean = 7.01, SD = 1.96),  $t(177) = 2.28$ ,  $p = .024$ . The effect size ( $d = .34$ ) again suggested a small difference.

- **Comparison of e-Argumentation Group and Software-free Argumentation Group:** There was no significant difference between the pre-test achievement scores of the two experimental groups,  $t(183) = -1.54$ ,  $p = .126$ . However, when comparing post-test achievement scores, the Software-free Argumentation group (Mean = 7.67, SD = 1.91) achieved significantly higher results than the e-Argumentation group (Mean = 6.71, SD = 1.74),  $t(183) = -3.55$ ,  $p < .001$ . The effect size for this difference ( $d = .52$ ) was moderate.

### **Findings related to the sixth grade level**

- **Comparison of e-Argumentation Group and Control Group:** No significant difference was observed between the groups in terms of pre-test achievement scores,  $t(74) = -0.29$ ,  $p = .776$ . When post-test scores were compared, the e-Argumentation group (Mean = 6.72, SD = 1.56) was found to be significantly more successful than the control group (Mean = 5.65, SD = 1.77),  $t(74) = 2.80$ ,  $p = .006$ . The effect size ( $d = .64$ ) indicated a moderate difference.
- **Comparison of the Software-Free Argumentation Group and the Control Group:** There was no significant difference between the groups regarding pre-test scores,  $t(80) = 0.84$ ,  $p = .401$ . In the post-test analysis, the Software-free Argumentation group (Mean = 6.05, SD = 1.19) showed significantly higher achievement compared to the control group (Mean = 5.37, SD = 1.45),  $t(80) = 2.31$ ,  $p = .024$ . The effect size ( $d = .52$ ) suggested a moderate difference.
- **Comparison of e-Argumentation Group and Software-free Argumentation Group:** The pre-test achievement scores of the two experimental groups were similar,  $t(76) = -0.48$ ,  $p = .635$ . Upon comparing the post-test achievement scores, the e-Argumentation group (Mean = 6.72, SD = 1.56) was found to be significantly more successful than the Software-free Argumentation group (Mean = 6.05, SD = 1.19),  $t(76) = 2.13$ ,  $p = .037$ . The calculated effect size ( $d = .48$ ) indicated a small-to-moderate difference.

### **Findings related to seventh grade level**

- **Comparison of e-Argumentation Group and Control Group:** No significant difference was observed between the pre-test achievement scores of the groups,  $t(122) = 1.33$ ,  $p = .185$ . Analysis of the post-test scores revealed that the e-Argumentation group (Mean = 10.77, SD = 2.16) achieved significantly higher results than the control group (Mean = 9.58, SD = 2.46),  $t(122) = 2.85$ ,  $p = .005$ . The effect size ( $d = .51$ ) corresponded to a moderate difference.
- **Comparison of Software-free Argumentation Group and Control Group:** No significant difference was found between these groups in terms of pre-test achievement scores,  $t(57) = -1.89$ ,  $p = .065$ . When post-test scores were examined, the Software-free Argumentation group (Mean = 11.14, SD = 1.87) was significantly more successful than the control group (Mean = 10.00, SD = 2.27),  $t(57) = 2.10$ ,  $p = .040$ . The effect size ( $d = .55$ ) indicated a moderate difference.

- **Comparison of e-Argumentation Group and Software-free Argumentation Group:** There was no significant difference between the pre-test scores of the two experimental groups,  $t(87) = 0.96$ ,  $p = .341$ . When post-test achievement scores were compared, no statistically significant difference was detected between the e-Argumentation group and the Software-free Argumentation group,  $t(87) = -0.79$ ,  $p = .430$ .

### ***Findings related to the eighth grade level***

- **Comparison of e-Argumentation Group and Control Group:** There was no significant difference between the groups in terms of pre-test achievement scores,  $t(74) = 0.96$ ,  $p = .338$ . Comparing the post-test scores, the e-Argumentation group (Mean = 10.00, SD = 1.89) demonstrated significantly higher achievement than the control group (Mean = 7.33, SD = 2.34),  $t(74) = 5.45$ ,  $p < .001$ . The effect size ( $d = 1.26$ ) was very large.
- **Comparison of the Software-Free Argumentation Group and the Control Group:** No significant difference was found between the groups regarding pre-test scores,  $t(59) = 0.07$ ,  $p = .945$ . In the post-test analysis, it was determined that the Software-free Argumentation group (Mean = 8.77, SD = 2.47) was significantly more successful than the control group (Mean = 7.52, SD = 2.13),  $t(59) = 2.12$ ,  $p = .038$ . The effect size ( $d = .54$ ) suggested a moderate difference.
- **Comparison of e-Argumentation Group and Software-free Argumentation Group:** The pre-test achievement scores of the two experimental groups were similar,  $t(65) = -0.52$ ,  $p = .607$ . When post-test achievement scores were compared, the e-Argumentation group (Mean = 10.00, SD = 1.89) was found to be significantly more successful than the Software-free Argumentation group (Mean = 8.77, SD = 2.47),  $t(65) = 2.32$ ,  $p = .024$ . The effect size for this difference ( $d = .56$ ) was moderate.

In general, these findings demonstrate that argumentation-based instruction (both with e-Argumentation support and without software) significantly enhanced students' science academic achievement across all grade levels compared to traditional methods (Bozkurt & Dođru, 2021; Kingir, 2011). The additional contribution of the e-Argumentation software was particularly evident in the 6th and 8th grades. However, the greater effectiveness of the non-software approach in the 5th grade and the absence of a significant difference in the 7th grade suggest that the impact of technology integration may vary depending on factors such as grade level.

### ***Findings related to argumentation skills***

The development of students' argumentation skills was assessed using rubrics appropriate to the argumentation approach employed. This assessment was based on the written arguments students constructed in science classes using various argumentation approaches (Argumentation-Based Science Learning [ABSL], Analytical Model, Socio-scientific Issues [SSI]), or on the data they entered into the e-Argumentation software. For the ABSL approach, a rubric adapted by Demirbađ and Günel (2014) from Choi (2008) was used, which evaluated argument quality across 39 items scored from 0–3

(total score range 0-117). For the Analytical Model and SSI approaches, categories adapted by Hiçde and Aktamış (2017) from Osborne et al. (2004) were utilised. These categories scored the basic components of an argument (claim, data, evidence, support, rebuttal) as 0 (absent), 1 (weak), or 2 (strong) in terms of quality and quantity, with a total score range of 0-10.

Due to technical limitations during the implementation (e.g. students in the e-Argumentation group sharing computers), the arguments of students in the e-Argumentation group were evaluated as group products rather than on an individual basis. This constraint precluded a direct, individual-level statistical comparison of argumentation skills between the e-Argumentation group and the Software-free Argumentation group. This methodological limitation has important implications for the interpretation of these findings. Specifically, the reliability of a group score as an indicator of individual student skill is uncertain, as it is unclear whether all students contributed equally to the group product or if the higher scores primarily reflect the performance of a few dominant members. Consequently, the results related to argumentation skills should be interpreted with caution. The findings can be generalised to the ability of groups to produce higher-quality arguments when using the software, but not necessarily to the argumentation skill development of every individual student within those groups.

As indicated in Table 4, across all grade levels (5, 6, 7, and 8) and all implemented argumentation approaches (ABSL/SWH, Analytical Model, SSI), the argument quality scores obtained by students in the e-Argumentation groups (as group averages) were generally observed to be higher than the average scores of students in the Software-free Argumentation groups. For instance, in the 5th-grade ABSL application, the

**Table 4.** Argumentation Quality Scores Table.

Class Level	Argumentation Approach	Group Type	Average Argument Quality Score
<b>Grade 5</b>	ABSL/SWH	E-Argumentation	43.80
		Software-Free Argumentation	30.00
	Analytical Model	E-Argumentation	3.42
		Software-Free Argumentation	2.21
	SSI	E-Argumentation	5.11
		Software-Free Argumentation	3.53
<b>Grade 6</b>	ABSL/SWH	E-Argumentation	44.80
		Software-Free Argumentation	25.05
	Analytical Model	E-Argumentation	3.27
		Software-Free Argumentation	2.12
	SSI	E-Argumentation	4.33
		Software-Free Argumentation	2.38
<b>Grade 7</b>	ABSL/SWH	E-Argumentation	53.00
		Software-Free Argumentation	34.79
	Analytical Model	E-Argumentation	3.67
		Software-Free Argumentation	1.90
	SSI	E-Argumentation	4.42
		Software-Free Argumentation	2.88
<b>Grade 8</b>	ABSL/SWH	E-Argumentation	46.00
		Software-Free Argumentation	29.09
	Analytical Model	E-Argumentation	5.42
		Software-Free Argumentation	2.88
	SSI	E-Argumentation	5.92
		Software-Free Argumentation	2.20

Note: Scores for the ABSL/SWH approach are between 0-117, while scores for the Analytic Model and SSI approaches are between 0-10. e-Scores in the Argumentation group are group averages.

average score for e-Argumentation groups was 43.80, whereas the average for the software-free argumentation groups was 30.00. Similarly, in the 8th-grade Analytical Model application, the e-Argumentation groups averaged 5.42, while the group without software averaged 2.88. This descriptive comparison suggests that the e-Argumentation software may offer a potential advantage over non-software applications in supporting students' abilities to formulate and present argument components such as claims, evidence, and justifications.

These observations are corroborated by students' statements regarding their own experiences. In interviews conducted as part of the qualitative dimension of the study, students in both the e-Argumentation and Software-free Argumentation groups generally expressed positive opinions about the argumentation process. Many students, particularly those in the e-Argumentation group, reported that during argumentation activities, they;

- Learned better and more permanently by discussing, considering different ideas, and defending their own.
- Found the lessons more enjoyable, a theme mentioned by 27 of the 80 students interviewed.
- Were more cognitively active during the process (i.e. they engaged in more thinking, questioning, and research).
- Experienced increased confidence in expressing their ideas, noting that even students who were typically quiet participated in the process.

These qualitative findings reinforce the impression that the argumentation process itself, and the e-Argumentation software when supported by technology, positively influenced not only students' academic achievement but also their engagement in the learning process and their perceived development of argumentation-related skills.

### ***Findings related to epistemological beliefs and reasoning***

Students' epistemological beliefs and reasoning were examined using three distinct data collection tools: (1) the Epistemological Belief Scale, measuring beliefs about the nature of knowledge and knowing; (2) the Epistemic Vignette, designed to identify epistemic positions (absolutist, multiplist, evaluativist); and (3) semi-structured interviews, aimed at uncovering students' perspectives on scientific knowledge following their argumentation experiences.

#### ***Epistemological belief scale results (Quantitative findings)***

Pre-test and post-test scores from the two sub-dimensions of the scale, 'nature of knowledge' and 'nature of knowing,' were compared between groups using multivariate analysis of variance/covariance (MANOVA/MANCOVA). For significant differences, eta-square ( $\eta^2$ ) effect sizes were calculated, with .01 interpreted as small, .06 as medium, and .14 as a large effect.

- **5th-Grade Level:** No significant difference was found between the pre-test scores of the groups (Wilks'  $\Lambda = .998$ ,  $p = .862$  for e-Argumentation vs. Control; Wilks'  $\Lambda$

= .999,  $p = .940$  for Software-free Argumentation vs. Control; Wilks'  $\Lambda = .996$ ,  $p = .684$  for e-Argumentation vs. Software-free Argumentation). When post-test scores were compared, no statistically significant differences in epistemological beliefs were found between the e-Argumentation group and the control group (Wilks'  $\Lambda = .992$ ,  $p = .494$ ), the Software-free Argumentation group and the control group (Wilks'  $\Lambda = .993$ ,  $p = .530$ ), or the e-Argumentation group and the Software-free Argumentation group (Wilks'  $\Lambda = .980$ ,  $p = .161$ ).

- **6th-Grade Level:** A significant difference was found between the pre-test scores of the e-Argumentation and control groups (Wilks'  $\Lambda = .892$ ,  $p = .015$ ); therefore, post-tests were compared using MANCOVA (controlling for pre-test scores), and no significant difference was found (Wilks'  $\Lambda = .998$ ,  $p = .915$ ). A significant difference was also found between the pre-test scores of the Software-free Argumentation and control groups (Wilks'  $\Lambda = .908$ ,  $p = .022$ ); MANCOVA revealed no significant difference between their post-test scores (Wilks'  $\Lambda = .969$ ,  $p = .301$ ). Similarly, a significant pre-test difference existed between the two experimental groups (e-Argumentation and Software-free Argumentation) (Wilks'  $\Lambda = .917$ ,  $p = .040$ ), but MANCOVA showed no significant difference in their post-test scores (Wilks'  $\Lambda = .984$ ,  $p = .548$ ).
- **7th-Grade Level:** No significant differences were found between the pre-test scores of the groups ( $p > .05$ ). When post-test scores were compared, the epistemological belief scores of the e-Argumentation group were found to be significantly higher than those of the control group (Wilks'  $\Lambda = .943$ ,  $p = .029$ ,  $\eta^2 = .057$ ). Analysis of variance results indicated that this difference was significant for both the knowledge sub-dimension ( $F(1, 122) = 4.74$ ,  $p = .031$ ,  $\eta^2 = .037$ ) and the knowing sub-dimension ( $F(1, 122) = 6.74$ ,  $p = .011$ ,  $\eta^2 = .052$ ). The effect sizes were generally small. However, no significant difference in post-test scores was detected between the Software-free Argumentation group and the control group (Wilks'  $\Lambda = .911$ ,  $p = .073$ ) or between the e-Argumentation group and the Software-free Argumentation group (Wilks'  $\Lambda = .985$ ,  $p = .521$ ).
- **8th-Grade Level:** No significant difference was found between the pre-test scores of the groups ( $p > .05$ ). Upon comparing post-test scores, the epistemological belief scores of the e-Argumentation group were significantly higher than those of the control group (Wilks'  $\Lambda = .864$ ,  $p = .005$ ,  $\eta^2 = .136$ ). This difference was significant for both the knowledge subscale ( $F(1, 74) = 11.17$ ,  $p = .001$ ,  $\eta^2 = .131$ ) and the knowing subscale ( $F(1, 74) = 7.12$ ,  $p = .009$ ,  $\eta^2 = .088$ ). The total effect size approached a moderate level. However, no significant difference in post-test scores was found between the Software-free Argumentation group and the control group (Wilks'  $\Lambda = .938$ ,  $p = .155$ ) or between the e-Argumentation group and the Software-free Argumentation group (Wilks'  $\Lambda = .977$ ,  $p = .470$ ).

In general, the findings from the Epistemological Belief Scale suggest that the e-Argumentation intervention had a statistically significant but small-to-moderate effect on improving students' epistemological beliefs compared to traditional teaching, particularly at the 7th and 8th-grade levels. However, Software-free Argumentation, or the comparison between the two different argumentation practices (e-Argumentation vs. Software-free Argumentation), did not yield statistically significant differences in the beliefs measured by this scale.

### ***Epistemic vignette results (Qualitative findings)***

The epistemic vignette was administered to the same sample of 80 students who participated in the semi-structured interviews. As previously described in the methodology, this sample was composed of at least two volunteers from each experimental classroom, ensuring that the responses represented a cross-section of students from all implementation contexts. The qualitative analysis of responses to the epistemic vignette (Kuhn et al., 2008), employed to determine students' epistemic positions, revealed a different pattern. Examination of the responses from students participating in the qualitative dimension of the study (N = 80) indicated that a substantial proportion of students across different groups (e-Argumentation, Software-free Argumentation, Control) and all grade levels (5-8) predominantly held an 'Absolutist' epistemological position, which accepts knowledge as certain, unchanging, and based on authority. Although a minority of students used expressions indicative of 'Multiplicist' or 'Evaluativist' positions, the prevailing profile was 'Absolutist.' It was concluded that the argumentation-based interventions (both e-Argumentation and Software-free Argumentation) did not lead to a discernible change in these fundamental epistemological positions as measured by the vignette analysis. This finding aligns with expectations from the literature regarding the cognitive developmental levels of middle school students (e.g. Kuhn et al., 2008).

### ***Student interviews: perspectives on scientific knowledge (Qualitative findings)***

Contrary to the findings from the Epistemological Belief Scale and the vignette, semi-structured interviews with students in the experimental groups suggested that the argumentation process positively influenced their perceptions and awareness of scientific knowledge. Many students in both the e-Argumentation and Software-free Argumentation groups reported that, at the end of this process, they;

- Began to question information more frequently, a sentiment expressed by 24 of the 80 participants.
- Realised that scientific knowledge might not be absolute and unchangeable, and that it could evolve with new evidence or be subject to different interpretations.
- Better understood the importance of evidence in verifying the truth of a claim.
- Understood that the opinions of authorities (scientists, teachers, books, etc.) or the majority could also be questioned and might not always represent absolute truth.
- Felt the need to check the reliability of information obtained from the internet or other sources and to compare different sources.
- Believed they should strive for greater objectivity and impartiality when making decisions about an issue or defending a claim

These statements suggest that the argumentation process, although challenging to measure precisely, stimulated students' epistemic awareness and triggered a shift in their approach to knowledge. The students' own accounts indicate a trend towards a more sophisticated epistemic understanding than what was captured by the scale and vignette results.

## Discussion

This study investigated the effects of argumentation instruction supported by e-Argumentation software on middle school students' science academic achievement, argumentation skills, and epistemological beliefs, comparing these to argumentation instruction without software support and traditional science instruction. The findings generally indicate that argumentation-based teaching yields positive effects on student outcomes, and that e-Argumentation software can effectively support the learning process.

### *Effects on academic achievement*

One of the most consistent findings of this study is that both e-Argumentation-supported and software-free argumentation practices significantly enhanced students' academic achievement in science compared to traditional teaching methods across all grade levels. This result aligns with previous literature indicating that argumentation deepens students' conceptual understanding and supports science learning (e.g. Asterhan & Schwarz, 2007; Newton et al., 1999; Zohar & Nemet, 2002). The argumentation process positions students as active learners; as they construct claims, evaluate evidence, and provide justifications, they engage with scientific concepts more profoundly (Driver et al., 2000; Osborne et al., 2004). This cognitive engagement likely contributed to more permanent knowledge retention and transferability to different situations, thereby positively influencing academic achievement.

Regarding the additional contribution of the e-Argumentation software, the significantly higher achievement of software-using groups compared to software-free argumentation groups at the 6th and 8th-grade levels suggests that technology support may have more effectively reinforced the learning processes of students at these grade levels. This could be attributed to the structured argument construction environment and the facilities for visualising and organising evidence provided by the software. Such scaffolding may have assisted students in managing complex information more effectively, consequently boosting their academic achievement. However, the finding that the non-software argumentation group outperformed the e-Argumentation group in the 5th grade is noteworthy. This outcome might be related to factors such as the adaptation processes of younger students to technology use, the perceived complexity of the interface, or their not-yet-fully-developed abstract thinking skills. In the 7th grade, the absence of a significant difference in academic achievement between the two experimental approaches suggests that students at this grade level may have benefited similarly from both types of argumentation, or that other variables influenced this result.

### *Effects on argumentation skills*

In terms of argumentation skills, the argument quality scores of students in the e-Argumentation groups (as group averages) were generally observed to be higher than the average scores of students in the Software-free Argumentation groups across all grade levels and in different argumentation approaches. This descriptive finding suggests that the e-Argumentation software can serve as an effective scaffold in developing

students' skills in constructing and presenting argument components such as claims, evidence, and justifications. The software's modules for visualising argument structures, adding diverse types of evidence (text, visuals, data tables, etc.), and organising them may have helped students overcome common challenges in the argumentation process (McNeill & Krajcik, 2008; Sandoval & Millwood, 2005). The qualitative data, where students reported finding the e-Argumentation process more engaging and feeling more active, may also have indirectly supported this skill development.

However, the evaluation of arguments in the e-Argumentation group as group products limits the ability to draw definitive conclusions about individual skill development. It remains unclear whether all students within a group contributed equally to the process, or whether the high scores observed reflect the performance of a few dominant students. This constitutes an important methodological limitation that should be considered when interpreting these findings.

### ***Effects on epistemological beliefs and reasoning***

The impact on epistemological beliefs and reasoning presents a more complex picture in this study. The results from the Epistemological Belief Scale indicated that the e-Argumentation intervention had a statistically significant but small-to-moderate effect on improving students' epistemological beliefs compared to traditional teaching, particularly at the 7th and 8th-grade levels. In most other cases, Software-free Argumentation, or a comparison between the two different argumentation practices, did not yield statistically significant differences in the beliefs measured by this scale. Similarly, Epistemic Vignette analyses revealed that the vast majority of students, irrespective of the intervention type, remained in the 'Absolutist' epistemological position, which is developmentally anticipated (Kuhn et al., 2008). These findings underscore the difficulty of altering epistemological beliefs through short-term interventions and suggest that such changes may necessitate longer-term and more profound epistemic practices (Hofer & Pintrich, 1997). Although argumentation practices introduce students to diverse perspectives and encourage evidence-based thinking, they alone may be insufficient to transform deeply ingrained epistemological assumptions.

Nevertheless, somewhat contradicting the quantitative findings, qualitative interviews with the students provided significant indications that the argumentation process (both e-Argumentation and without software) positively influenced their perceptions and awareness of scientific knowledge. Students' statements about questioning knowledge more, grasping the provisional nature of knowledge, and understanding the importance of evidence suggest that argumentation can indeed trigger epistemic awareness. This might signify the nascent stages of an epistemic development not fully captured by the scales or vignettes, or it could imply that the impact of argumentation is more pronounced on epistemic practices and discourse than on epistemic beliefs per se (Sandoval & Millwood, 2007). The limited advantage shown by the e-Argumentation group over the control group in the upper grades might be a consequence of encountering a greater number or variety of evidence and opinions facilitated by the software.

A more nuanced interpretation of the findings must also consider key contextual factors. For instance, variability in teacher implementation fidelity and the suitability

of the argumentation approach for different science units likely influenced the mixed outcomes across grade levels. Furthermore, the limited impact on epistemological beliefs can be attributed to the short, one-semester intervention duration and a lack of tasks designed for sustained epistemic reflection. These insights translate into practical recommendations, highlighting the need for improved school technical infrastructure and, crucially, long-term professional development that provides sustained, on-the-job support for teachers.

### ***Limitations of the study***

It is important to consider certain limitations when interpreting the findings of this study. Firstly, the quasi-experimental design employed made it challenging to achieve full control over all variables and did not entirely preclude the possibility that pre-existing, unobserved differences between groups might have influenced the results. Secondly, the assessment of argumentation skills in the e-Argumentation group at the group level prevented drawing firm conclusions about individual development. Thirdly, the study was conducted in schools within a specific geographical region, which may limit the generalizability of the findings to different socio-cultural contexts. Fourthly, potential variations in how teachers implemented the argumentation approaches could have affected student outcomes; this factor was not fully controlled. Finally, measuring change in epistemological beliefs is an inherently complex process, and the measurement tools used may not have captured all dimensions of this change.

### **Conclusion and recommendations**

This study examined the effects of argumentation instruction, with and without the support of e-Argumentation software, on middle school students' science academic achievement, argumentation skills, and epistemological beliefs. The main conclusions drawn from the findings, and recommendations based thereon, are presented below, categorised by the relevant variables.

#### ***Academic achievement***

The findings demonstrated that argumentation-based instruction (both with e-Argumentation support and without software) significantly enhanced students' academic achievement in science compared to traditional teaching methods across all grade levels (grades 5-8). Groups utilising the e-Argumentation software exhibited a significant advantage in academic achievement at the 6th and 8th-grade levels when compared to argumentation groups without software. However, at the 5th-grade level, Software-free Argumentation resulted in higher academic achievement than e-Argumentation, while at the 7th-grade level, no significant difference in academic achievement was observed between the two experimental approaches. These findings underscore the importance of more widespread adoption of argumentation-based teaching approaches to improve students' academic achievement in science courses. When employing technology-supported tools, it should be recognised that the software's impact may vary depending on factors such as grade level and student readiness; therefore, implementation should

be tailored to the specific characteristics of the student group. Future research that comparatively examines the responses and learning outcomes of students at different grade levels or in diverse socio-cultural contexts to technology-supported argumentation practices would make a valuable contribution to the field.

### ***Argumentation skills***

Descriptively, the argument quality scores of students in the e-Argumentation groups (as group averages) were generally observed to be higher than the average scores of students in the Software-free Argumentation groups across all grade levels and in different argumentation approaches. This indicates that e-Argumentation software holds potential as a tool for supporting argument construction and presentation skills. Qualitative interviews with students also corroborated the impression that the argumentation process, and particularly the e-Argumentation software, positively influenced engagement in the learning process and skill development. To leverage this potential more effectively, the value of technology-supported tools like e-Argumentation should be utilised, especially in bolstering students' abilities to construct arguments, present evidence, and interact with peers. Teachers should be encouraged to integrate such software into their lessons to develop students' argumentation skills. However, it is crucial to specifically address areas where students typically struggle, such as using evidence and reasoning when constructing arguments, and to provide additional support or scaffolds to develop these skills. Considering the limitation of assessing argumentation skills at the group level in the e-Argumentation group, it is imperative to provide the necessary technical infrastructure (e.g. a device for each student) and to ensure that software includes features allowing for individual assessment to better track students' individual contributions and progress in technology-supported argumentation practices. Future research could delve into the specific effects of different features of the e-Argumentation software (e.g. visualisation tools, feedback mechanisms, collaborative modules) on students' distinct skills (e.g. evidence evaluation, rebuttal, consensus-building) and could focus on research designs and implementation conditions that allow for individual-level assessment of argumentation skills in the e-Argumentation group.

### ***Epistemological beliefs and reasoning***

The analysis of the impact on epistemological beliefs and reasoning revealed that the effect of argumentation-based interventions on students' epistemological beliefs was more limited. According to the Epistemological Belief Scale results, the e-Argumentation intervention significantly improved students' epistemological beliefs, albeit to a small-to-moderate extent, compared to traditional teaching, but only at the 7th and 8th-grade levels. In other instances, no statistically significant differences were found. Similarly, Epistemic Vignette analyses indicated that the majority of students remained in the 'Absolutist' epistemological position, irrespective of the intervention. Nevertheless, qualitative interviews with students suggested that the argumentation process positively influenced their perceptions and awareness of scientific knowledge (e.g. questioning knowledge, the importance of evidence, the provisional nature of knowledge) and could trigger an epistemic awakening. While the argumentation process is thought to

have the potential to trigger epistemic awareness, merely conducting argumentation activities may not suffice for more fundamental shifts in epistemological beliefs. Therefore, if the aim is to cultivate students' epistemological beliefs, it is recommended to integrate explicit and reflective discussions or activities on topics such as the nature of knowledge and how scientific knowledge is acquired and evaluated, in addition to the argumentation process itself. Encouraging students to reflect on their own thinking processes and approaches to knowledge may be beneficial. Given that changes in epistemological beliefs require a sustained process, longitudinal studies that track the development of students' argumentation skills, and particularly their epistemological beliefs, over an extended period would be valuable for understanding the lasting effects of such interventions. Furthermore, a more comprehensive understanding could be achieved through mixed-design studies that employ a combination of quantitative data collection tools and vignettes, alongside diverse methods such as interviews, think-aloud protocols, or portfolios, to measure students' epistemological beliefs and reasoning.

Overall, this research robustly supports the notion that argumentation-based instruction in middle school science education is an effective approach for enhancing students' academic achievement and improving their argumentation skills. The e-Argumentation software demonstrates considerable potential in supporting these processes. The more complex and limited impact on epistemological beliefs suggests that fostering changes in such beliefs necessitates more long-term and multifaceted interventions. Finally, research into the professional development needs of teachers to effectively implement technology-supported argumentation practices in their classrooms, and the development of effective support models for these needs, will also significantly contribute to the field.

## **Acknowledgements**

The authors would like to thank the Department of Information Technologies of Bolu Abant İzzet Baysal University for providing hosting services for the web-based application developed within the scope of the project. We also extend our gratitude to the science teachers working in public middle schools in Bolu province who voluntarily participated in this project.

## **Disclosure statement**

No potential conflict of interest was reported by the author(s).

## **Funding**

This research was supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) under grant number 219K028. The e-Argumentation software used in this study was developed as part of this project. No other funding sources were involved in this research.

## **Ethics approval and consent to participate**

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Bolu Abant İzzet Baysal University Social Sciences Human Research Ethics Committee (Protocol No: 2018/183) and the Bolu Provincial Directorate of National Education (Approval Ref: 18291087, Date: 22.12.2020). Informed consent was obtained from all individual

participants (and their legal guardians, as participants were middle school students) included in the study.

## Availability of data and materials

The web-based application developed within the scope of the project can be accessed at <https://eargumantasyon.ibu.edu.tr/>. The research data will be shared upon reasonable request to the corresponding author.

## ORCID

Eralp Bahçivan  <http://orcid.org/0000-0001-5621-3302>

Nuh Yavuzalp  <http://orcid.org/0000-0001-9275-275X>

Mehmet Demirbağ  <http://orcid.org/0000-0002-0881-8486>

Erhan Güneş  <http://orcid.org/0000-0002-4268-4645>

Mutlu Tahsin Üstündağ  <http://orcid.org/0000-0001-6198-2819>

Niyazi Öztürk  <http://orcid.org/0000-0002-0637-6063>

## References

- Akpınar, Y., Ardac, D., & Amuce, N. E. (2015). Computer versus computer and human support in an argumentation-based science learning environment. *Journal of Online Learning Research*, 1(2), 137–161.
- Ansell, S., & Koenig, J. (2011). Cybertracker: An integral management tool used by rangers in the Djelk indigenous protected area, Central Arnhem land, Australia. *Ecological Management & Restoration*, 12(1), 13–25. <https://doi.org/10.1111/j.1442-8903.2011.00575.x>
- Asterhan, C. S., & Schwarz, B. B. (2007). The effects of monological and dialogical argumentation on concept learning in evolutionary theory. *Journal of Educational Psychology*, 99(3), 626–639. <https://doi.org/10.1037/0022-0663.99.3.626>
- Asterhan, C. S., & Schwarz, B. B. (2009). Transformation of robust misconceptions through peer argumentation. In C. S. Asterhan, & B. B. Schwarz (Eds.), *Transformation of knowledge through classroom interaction* (pp. 159–172). Routledge.
- Bahçivan, E. (2014). Examining relationships among Turkish preservice science teachers' conceptions of teaching and learning, scientific epistemological beliefs and science teaching efficacy beliefs. *Journal of Baltic Science Education*, 13(6), 870–882. <https://doi.org/10.33225/jbse/14.13.870>
- Bahçivan, E., & Cobern, W. W. (2016). Investigating coherence among Turkish elementary science teachers' teaching belief systems, pedagogical content knowledge and practice. *Australian Journal of Teacher Education*, 41(10), 63–80. <https://doi.org/10.14221/ajte.2016v41n10.5>
- Belland, B. R. (2010). Portraits of middle school students constructing evidence-based arguments during problem-based learning: The impact of computer-based scaffolds. *Educational Technology Research and Development*, 58(3), 285–309. <https://doi.org/10.1007/s11423-009-9139-4>
- Bozkurt, O., & Dođru, S. (2021). The effect of argumentation-based classroom activities on the academic achievement, logical thinking skills and willingness to discuss of fifth grade middle school students. *Mehmet Akif Ersoy University Journal of Faculty of Education*, 58, 624–644. <https://doi.org/10.21764/mauefd.775143>
- Bricker, L. A., & Bell, P. (2008). Conceptualizations of argumentation from science studies and the learning sciences and their implications for the practices of science education. *Science Education*, 92(3), 473–498. <https://doi.org/10.1002/sce.20278>
- Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K-12 science contexts. *Review of Educational Research*, 80(3), 336–371. <https://doi.org/10.3102/0034654310376953>

- Chin, C., & Osborne, J. (2010). Supporting argumentation through students' questions: Case studies in science classrooms. *The Journal of the Learning Sciences*, 19(2), 230–284. <https://doi.org/10.1080/10508400903530036>
- Choi, A. (2008). *A study of student written argument using the science writing heuristic approach in inquiry-based freshman general chemistry laboratory classes* [Unpublished doctoral dissertation]. University of Iowa.
- Conley, A. M., Pintrich, P. R., Vekiri, L., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 186–204. <https://doi.org/10.1016/j.cedpsych.2004.01.004>
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Sage.
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research*. Sage publications.
- Delen, I. (2014). *Supporting students' scientific explanations: A case study investigating the synergy focusing on a teacher's practices when providing instruction and using mobile devices* [Doctoral Dissertation]. Michigan State University.
- Demirbağ, M., & Bahçıvan, E. (2022). Psychological modeling of preservice science teachers' argumentativeness, achievement goals, and epistemological beliefs: A mixed design. *European Journal of Psychology of Education*, 37(1), 257–278. <https://doi.org/10.1007/s10212-021-00558-w>
- Demirbağ, M., & Günel, M. (2014). Integrating argument-based science inquiry with modal representations: Impact on science achievement, argumentation, and writing skills. *Educational Sciences: Theory & Practice*, 14(1), 386–391. <https://doi.org/10.12738/estp.2014.1.1994>
- Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, 84(3), 287–312. [https://doi.org/10.1002/\(SICI\)1098-237X\(200005\)84:3<287::AID-SCE1>3.0.CO;2-A](https://doi.org/10.1002/(SICI)1098-237X(200005)84:3<287::AID-SCE1>3.0.CO;2-A)
- Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38(1), 39–72. <https://doi.org/10.1080/03057260208560187>
- Erduran, S., & Jimenez-Aleixandre, M. P. (Eds.). (2007). *Argumentation in science education: Recent developments and future directions*. Springer.
- Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's Argument Pattern for studying science discourse. *Science Education*, 88(6), 915–933. <https://doi.org/10.1002/sce.20012>
- Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, 48(1), 56–66. <https://doi.org/10.1080/00461520.2012.748005>
- Green, S. B., & Salkind, N. J. (2014). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (7th ed.). Pearson.
- Hiğde, E., & Aktamış, H. (2017). Reflection of explicit-reflective argumentation based and explicit-reflective nature of science teaching on prospective science teachers' written arguments. *Çukurova University Faculty of Education Journal*, 46(1), 39–84. <https://doi.org/10.14812/uefd.309431>
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140. <https://doi.org/10.3102/00346543067001088>
- Kılıç, M. (2020). *Examining the decision-making processes of gifted students and their parents on a controversial text in the context of epistemic profiles: Bursa BİLSEM sample* [Master's Thesis]. Bursa Uludağ University.
- Kingır, S. (2011). *The use of argumentation based science learning approach in enabling students to understand the concepts of chemical change and mixture* [Doctoral Dissertation]. Middle East Technical University.
- Krajcik, J., Blumenfeld, P. C., Marx, R. W., Bass, K. M., & Fredericks, J. (1998). Inquiry in project-based science classrooms: Initial attempts by middle school students. *Journal of the Learning Sciences*, 7(3-4), 313–350. <https://doi.org/10.1080/10508406.1998.9672057>

- Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30(3), 411–433. <https://doi.org/10.1111/j.1468-2958.2004.tb00738.x>
- Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.
- Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319–337. <https://doi.org/10.1002/sce.3730770306>
- Kuhn, D. (2005). *Education for thinking*. Harvard University Press.
- Kuhn, D., Jordanou, K., Pease, M., & Wirkala, C. (2008). Beyond control of variables: What needs to develop to achieve skilled scientific thinking? *Cognitive Development*, 23(4), 435–451. <https://doi.org/10.1016/j.cogdev.2008.09.006>
- Kuhn, D., & Udell, W. (2003). The development of argument skills. *Child Development*, 74(5), 1245–1260. <https://doi.org/10.1111/1467-8624.00605>
- Laru, J., Järvelä, S., & Clariana, R. (2012). Supporting collaborative inquiry during a biology field trip with mobile peer-to-peer tools for learning: A case study with K-12 learners. *Interactive Learning Environments*, 20(2), 103–117. <https://doi.org/10.1080/10494821003771350>
- Mason, L., & Scirica, F. (2006). Prediction of students' argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction*, 16(5), 492–509. <https://doi.org/10.1016/j.learninstruc.2006.09.007>
- McNeill, K. L., & Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In M. C. Lovett, & P. Shah (Eds.), *Thinking with data: 33rd Carnegie Symposium on cognition* (pp. 233–265). Psychology Press.
- McNeill, K. L., & Krajcik, J. (2008). Inquiry and scientific explanations: Helping students use evidence and reasoning. In J. Luft, R. L. Bell, & J. Gess-Newsome (Eds.), *Science as inquiry in the secondary setting* (pp. 121–134). NSTA Press.
- Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education*, 21(5), 553–576. <https://doi.org/10.1080/095006999290570>
- Nussbaum, E. M., & Bendixen, L. D. (2003). Approaching and avoiding arguments: The role of epistemological beliefs, need for cognition, and extraverted personality traits. *Contemporary Educational Psychology*, 28(4), 573–595. [https://doi.org/10.1016/S0361-476X\(02\)00062-0](https://doi.org/10.1016/S0361-476X(02)00062-0)
- Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argumentation in school science. *Journal of Research in Science Teaching*, 41(10), 994–1020. <https://doi.org/10.1002/tea.20035>
- Pallant, J. (2005). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows* (2nd ed.). Open University Press.
- Sampson, V., & Clark, D. (2006). Assessment of argument in science education: A critical review of the literature. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the 7th International Conference of the Learning Sciences: Indiana University* (Vol. 2, pp. 655–661). International Society of the Learning Sciences.
- Sandoval, W. A., & Millwood, K. A. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction*, 23(1), 23–55. [https://doi.org/10.1207/s1532690xci2301\\_2](https://doi.org/10.1207/s1532690xci2301_2)
- Sandoval, W. A., & Millwood, K. A. (2007). What can argumentation tell us about epistemology? In S. Erduran, & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education* (pp. 71–88). Springer.
- Sandoval, W. A., & Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic supports for science inquiry. *Science Education*, 88(3), 345–372. <https://doi.org/10.1002/sce.10130>
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3), 498–504. <https://doi.org/10.1037/0022-0663.82.3.498>
- Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education*, 28(2-3), 235–260. <https://doi.org/10.1080/09500690500336957>
- Toulmin, S. E. (1958). *The uses of argument*. Cambridge University Press.

- Tsai, C. Y., Jack, B. M., Huang, T. C., & Yang, J. T. (2012). Using the cognitive apprenticeship web-based argumentation system to improve argumentation instruction. *Journal of Science Education and Technology*, 21(4), 476–486. <https://doi.org/10.1007/s10956-011-9339-7>
- Van Eemeren, F. H., & Grootendorst, R. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge University Press.
- Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35–62. <https://doi.org/10.1002/tea.10008>