Science & Education (2023) 32:1165–1195
https://doi.org/10.1007/s11191-022-00369-5

ARTICLE

Developing Students' Critical Thinking Skills and Argumentation Abilities Through Augmented Reality–Based Argumentation Activities in Science Classes

Tuba Demircioglu[1] · Memet Karakus[2] · Sedat Ucar[1]

Accepted: 15 July 2022 / Published online: 22 August 2022

Abstract
Due to the COVID-19 pandemic and adapting the classes urgently to distance learning, directing students' interest in the course content became challenging. The solution to this challenge emerges through creative pedagogies that integrate the instructional methods with new technologies like augmented reality (AR). Although the use of AR in science education is increasing, the integration of AR into science classes is still naive. The lack of the ability to identify misinformation in the COVID-19 pandemic process has revealed the importance of developing students' critical thinking skills and argumentation abilities. The purpose of this study was to examine the change in critical thinking skills and argumentation abilities through augmented reality–based argumentation activities in teaching astronomy content. The participants were 79 seventh grade students from a private school. In this case study, the examination of the verbal arguments of students showed that all groups engaged in the argumentation and produced quality arguments. The critical thinking skills of the students developed until the middle of the intervention, and the frequency of using critical thinking skills varied after the middle of the intervention. The findings highlight the role of AR-based argumentation activities in students' critical thinking skills and argumentation in science education.

1 Introduction

With rapidly developing technology, the number of children using mobile handheld devices has increased drastically (Rideout et al., 2010; Squire, 2006). Technologies and digital enhancements that use the internet have become a part of the daily life of school-age children (Kennedy et al., 2008), and education evolves in line with the changing

✉ Tuba Demircioglu
tubademircioglu@gmail.com

Memet Karakus
memkks@gmail.com

Sedat Ucar
sedatucar@gmail.com

1 Faculty of Education, Department of Mathematics and Science Education/Elementary Science Education, Cukurova University, 01330 Saricam-Adana, Turkey

2 Department of Educational Sciences, Cukurova University, Adana, Turkey

technology. Rapidly changing innovation technologies have changed the characteristics of learners in the fields of knowledge, skills, and expertise that are valuable for society, and circumstances for teachers and students have changed over time (Yuen et al., 2011). Almost every school subject incorporates technological devices into the pedagogy to different extents, but science teachers are the most eager to use technological devices in science classes because of the nature of the content they are expected to teach.

The COVID-19 pandemic has had an important impact on educational systems worldwide. Due to the fast-spreading of that disease, the educators had to adapt their classes urgently to technology and distance learning (Dietrich et al., 2020), and schools have had to put more effort into adapting new technologies to teaching. Z generation was born into a time of information technology, but they did not choose distance courses that were not created for them so they are not motivated during the classes (Dietrich et al., 2020). Directing students' interest in the course content is challenging, while their interest has changed by this technological development. The solution to this challenge emerges through creative pedagogies that integrate the instructional methods with new striking technology. Augmented reality has demonstrated high potential as part of many teaching methods.

2 Literature Review

2.1 Augmented Reality, Education, and Science Education

AR applications have important potential for many areas where rapid transfer of information is important. This is especially effective for education. Science education is among the disciplines where rapid information transfer is important. Taylor (1987, p. 1) stated that "the transfer of scientific and technological information to children and to the general public is as important as the search for information." With the rapid change in science and technology and outdating of knowledge, learning needs rapid changes in transfer of information (Ploman, 1987). Technology provides new and innovative methods for science education and could be an effective media in promoting students' learning (Virata & Castro, 2019). AR technology could be a promising teaching tool for science teaching in which AR technology is especially applicable (Arici et al., 2019).

Research shows that AR has great potential and benefits for learning and teaching (Yuen et al., 2011). The AR applications used in teaching and learning present many objects, practices, and experiments that students cannot obtain from the first-hand experience into many different dimensions because of the impossibilities in the real world, and it is an approach that can be applied to many science contents that are unreachable, unobtrusive, and unable to travel (Cai et al., 2013; Huang et al., 2019; Pellas et al., 2019). For example, physically unreachable phenomena such as solar systems, moon phases, and magnetic fields become accessible for learners through AR (Fleck & Simon, 2013; Kerawalla et al., 2006; Shelton & Hedley, 2002; Sin & Zaman, 2010; Yen et al., 2013). Through AR, learners can obtain instant access to location-specific information provided by a wide range of sources (Yuen et al., 2011). Location-based information, when used in particular contextual learning activities, is essential for assisting students' outdoor learning. This interaction develops comprehension, understanding, imagination, and retention, which are the learning and cognitive skills of learners (Chiang et al., 2014). For example, an AR-based mobile learning system was used in the study conducted by Chiang et al. (2014) on aquatic animals and plants. The location module can identify the students' GPS location, direct them to discover the target ecological regions, and provide the appropriate learning tasks or additional resources. When students explore various characteristics of learning objects, the camera and image editing modules can take the image from the real environment and make comment on the image of the observed things.

Research reveals that the use of AR technology as part of teaching a subject has the features of being constructivist, problem solving-based, student-centered, authentic, participative, creative, personalized, meaningful, challenging, collaborative, interactive, entertaining, cognitively rich, contextual, and motivational (Dunleavy et al., 2009). Despite its advantages and although the use of AR in science education is increasing, the integration of AR into science classes is still naive, and teachers still do not consider themselves as ready for use of AR in their class (Oleksiuk & Oleksiuk, 2020; Romano et al., 2020) and choose not to use AR technology (Alalwan et al., 2020; Garzón et al., 2019), because most of them do not have the abilities and motivation to design AR learning practices (Garzón et al., 2019; Romano et al., 2020). It is thought that the current study will contribute to the use of AR in science lessons and how science teachers will include AR technology in their lessons.

2.2 Argumentation, Critical Thinking, and Augmented Reality

New trends in information technologies have contributed to the development of new skills in which people have to struggle with a range of information and evaluate this information. An important point of these skills is the ability to argue with evidence (Jiménez -Aleixandre & Erduran, 2007) in which young people create appropriate results from the information and evidence given to them to criticize the claims of others in the direction of the evidence and to distinguish an idea from evidence-based situations (OECD, 2003, p. 132).

Learning with technologies could produce information and misinformation simultaneously (Chai et al., 2015). Misinformation has spread very quickly in public in COVID-19 pandemic, so the lack of the ability to interpret and evaluate the validity and credibility of them arose again (Saribas & Çetinkaya, 2021). This process revealed the importance of developing students' critical thinking skills and argumentation abilities (Erduran, 2020) to make decisions and adequate judgments when they encountered contradicting information (Chai et al., 2015).

Thinking about different subjects, evaluating the validity of scientific claims, and interpreting and evaluating evidence are important elements of science courses and play important roles in the construction of scientific knowledge (Driver et al., 2000). The use of scientific knowledge in everyday life ensures that critical thinking skills come to the forefront. Ennis (2011, p. 1) defined critical thinking as "Critical thinking is reasonable and reflective thinking focused on deciding what to believe". Jiménez-Aleixandre and Puig (2012) found this definition very broad, and they proposed a comprehensive definition of critical thinking that combines the components of social emancipation and evidence evaluation. It contains the competence to form autonomous ideas as well as the ability to participate in and reflect on the world around us. Figure 1 summarizes this comprehensive definition.

Critical thinking skills that include the ability to evaluate arguments and counterarguments in a variety of contexts are very important, and effective argumentation is

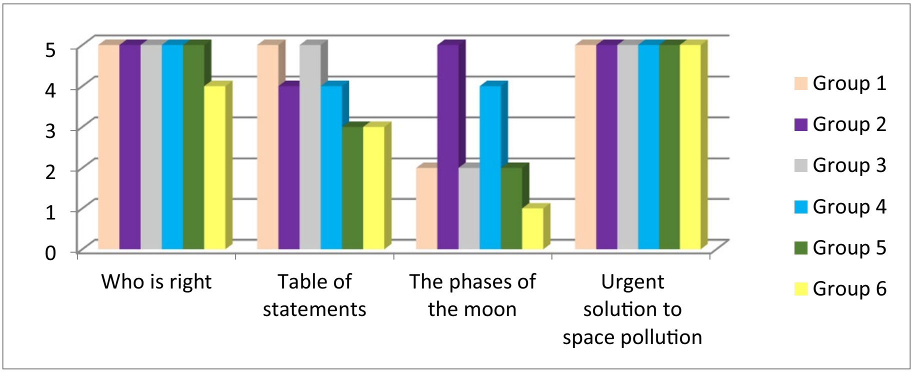

Fig. 1 Argumentation levels by groups

the focal point of criticism and the informed decision (Nussbaum, 2008). Argumentation is defined as the process of making claims about a scientific subject, supporting them with data, using warrants, and criticizing, refuting, and evaluating an idea (Toulmin, 1990). Argumentation as an instructional method is an important research area in science education and has received enduring interest from science educators for more than a decade (Erduran et al., 2015). Researchers concluded that learners mostly made only claims in the argumentation process and had difficulty producing well-justified and high-quality arguments (Demircioglu & Ucar, 2014; Demircioglu & Ucar, 2015; Cavagnetto et al., 2010; Erdogan et al., 2017; Erduran et al., 2004; Novak & Treagust, 2017). To improve the quality of arguments, students should be given supportive elements to produce more consistent arguments during argumentation. One of these supportive elements is the visual representations of the phenomena.

Visual representations could make it easier to see the structure of the arguments of learners (Akpınar et al., 2014) and improve students' awareness. For example, the number of words and comments used by students or meaningful links in conversations increases with visually enriched arguments (Erkens & Janssen, 2006). Sandoval & Millwood (2005) stated that students should be able to evaluate different kinds of evidence such as digital data and graphic photography to defend their claims. Appropriate data can directly support a claim and allow an argument to be accepted or rejected by students (Lin & Mintzes, 2010). Enriched visual representations provide students with detailed and meaningful information about the subject (Clark et al., 2007). Students collect evidence for argumentation by observing enriched representations (Clark et al., 2007), and these representations help to construct higher-quality arguments (Buckingham Shum et al., 1997; Jermann & Dillenbourg, 2003). Visualization techniques enable students to observe how objects behave and interact and provide an easy-to-understand presentation of scientific facts that are difficult to understand with textual or oral explanations (Cadmus, 1990). In short, technological opportunities to create visually enriched representations increase students' access to rich data to support their arguments.

Among the many technological opportunities to promote argumentation, AR seems to be the most promising application for instructing school subjects. AR applications are concerned with the combination of computer-generated data (virtual reality) and the real world, where computer graphics are projected onto real-time video images (Dias, 2009). In addition, augmented reality provides users with the ability to see a real-world environment enriched with 3D images and to interact in real time by combining virtual objects with the real environment in 3D and showing the spatial relations (Kerawalla et al., 2006). AR applications are thus important tools for students' arguments with the help of detailed and meaningful information and enriched representations. Research studies using AR technology revealed that all students in the study engaged in argumentation and produced arguments (Jan, 2009; Squire & Jan, 2007).

Many studies focusing on using AR in science education have been published in recent decades. Research studies related to AR in science education have focused on the use of game-based AR in science education (Atwood-Blaine & Huffman, 2017; Bressler & Bodzin, 2013; Dunleavy et al., 2009; López-Faican & Jaen, 2020; Squire, 2006), academic achievement (Hsiao et al., 2016; Faridi et al., 2020; Hwang et al., 2016; Lu et al., 2020; Sahin & Yilmaz, 2020;, Yildirim & Seckin-Kapucu, 2020), understanding science content and its conceptual understanding (Cai et al., 2021; Chang et al., 2013; Chen & Liu, 2020; Ibáñez et al., 2014), attitude (Sahin & Yilmaz, 20200; Hwang et al., 2016), self-efficacy (Cai et al., 2021), motivation (Bressler & Bodzin, 2013; Chen & Liu, 2020; Kirikkaya & Başgül, 2019; Lu et al., 2020; Zhang et al., 2014), and critical thinking skills (Faridi et al., 2020; Syawaludin et al., 2019). The general trend in these research studies based on the content of "learning/academic achievement," "understanding science content and its conceptual understanding," "motivation," "attitude," and methodologically quantitative studies was mostly used in articles in science education. Therefore, qualitative and quantitative data to be obtained from studies investigating the use of augmented reality technology in education and focusing on cognitive issues, interaction, and collaborative activities are needed (Arici et al., 2019; Cheng & Tsai, 2013).

Instructional strategies using AR technology ensure interactions between students and additionally between students and teachers (Hanid et al., 2020). Both the technological features of AR and learning strategies should be regarded by the teachers, the curriculum, and AR technology developers to acquire the complete advantage of AR in student learning (Garzón & Acevedo, 2019; Garzón et al., 2020). Researchers investigated the learning outcomes with AR-integrated learning strategies such as collaborative learning (Baran et al., 2020; Chen & Liu, 2020; Ke & Carafano, 2016), socioscientific reasoning (Chang et al., 2020), student-centered hands-on learning activities (Chen & Liu, 2020), inquiry-based learning (Radu & Schneider, 2019), concept-map learning system (Chen et al., 2019), problem-based learning (Fidan & Tuncel, 2019), and argumentation (Jan, 2009; Squire & Jan, 2007) in science learning.

The only two existing studies using both AR and argumentation (Jan, 2009; Squire & Jan, 2007) focus on environmental education and use location-based augmented reality games through mobile devices to engage students in scientific argumentation. Studies combining AR and argumentation in astronomy education have not been found in the literature. In the current study, AR was integrated with argumentation in teaching astronomy content.

Studies have revealed that many topics in astronomy are very difficult to learn and that students have incorrect and naive concepts (Yu & Sahami, 2007). Many topics include three-dimensional (3D) spatial relationships between astronomical objects (Aktamış & Arıcı, 2013; Yu & Sahami, 2007). However, most of the traditional teaching materials used in astronomy education are two-dimensional (Aktamış & Arıcı, 2013). Teaching astronomy through photographs and 2D animations is not sufficient to understand the difficult and complex concepts of astronomy (Chen et al., 2007). Static visualization tools such as texts, photographs, and 3D models do not change over time and do not have continuous movement, while dynamic visualization tools such as videos or animations show continuous movement and change over time (Schnotz & Lowe, 2008). However, animation is the presentation of images on a computer screen (Rieber & Kini, 1991), not in the real world, and the users do not have a chance to manipulate the images (Setozaki et al., 2017). As a solution to this shortcoming, using 3D technology in science classes, especially AR technology for abstract concepts, has become a necessity (Sahin & Yilmaz, 2020). By facilitating interaction with real and virtual environment and supporting object manipulation, AR is possible to enhance educational benefits (Billinghurst, 2002). The students are not passive participants while using AR technology. For example, the animated 3D sun and Earth models are moved on a handheld platform that adjusts its orientation in accordance with the student's point of view in Shelton's study (2002). They found that the ability of students to manage "how" and "when" they are allowed to manipulate virtual 3D objects has a direct impact on learning complex spatial phenomena. Experimental results show that compared with traditional video teaching, AR multimedia video teaching method significantly improves students' learning (Chen et al., 2022).

This study, which integrates argumentation with new striking technology "AR" in astronomy education, clarifies the relationship between them and examines variables such as critical thinking skills and argumentation abilities that are essential in the era we live, making this research important.

2.3 Research Questions

The purpose of this study was to identify the change in critical thinking skills and argumentation abilities through augmented reality–based argumentation activities in teaching astronomy content. The following research questions guided this study:

RQ1: How do the critical thinking skills of students who participated in both augmented reality and argumentation activities on astronomy change during the study?

RQ2: How do the argumentation abilities of students who participated in both augmented reality and argumentation activities on astronomy change during the study?

3 Method

In this case study, we investigated the change of critical thinking skills and argumentation abilities of middle school students. Before the main intervention, a pilot study was conducted to observe the effectiveness of the prepared lesson plans in practice and to identify the problems in the implementation process. The pilot study was recorded with a camera. The camera recordings were watched by the researcher, and the difficulties in the implementation process were identified. In the main intervention, preventions were taken to overcome these difficulties. Table 1 illustrates that the problems encountered during the pilot study and the preventions taken to eliminate these problems.

Table 1 The solutions to the problems in the pilot study

Problems in the pilot study	Solutions to the problems in the main intervention
The students were asked to download the AR applications on their tablets before the pilot study. However, some students could not download the applications so they could not use some of them	In the main intervention, a suitable hour for the students was determined, and an internet connection was established in a classroom of the school. All AR applications were downloaded to the tablets with the students. Also, the researcher gave practical information to the students about how to use the applications and gave them the opportunity to use them as well. In this way, the students had an experience with the applications before the main intervention
Some students tried to detect markers with the cameras of their tablets without opening the AR application in the activities. Markers could not be detected because the program was not run	The activities were performed after all students opened the applications
Due to the long duration of the activities, too many activities in one lesson, and problems with AR applications, the pilot implementation period took longer than planned	Some of the activities were not included in the main intervention. The long duration of the activities was due to the problems experienced in AR applications. For this reason, the above-mentioned solutions were implemented during the main intervention. Students were given a certain amount of time to do the activities

During the main intervention, qualitative data were collected through observations and audio recordings to determine the change in the critical thinking skills and argumentation abilities of students who participated in both augmented reality and argumentation activities on astronomy.

3.1 Context and Participants

The participants consisted of 79 7th middle school students aged between 12 and 13 from a private school in Southern Turkey. The participants were determined as students in a private school where tablet computers are available for each student and the school willing to participate in the study. Twenty-six students, including 17 females and 9 males, participated in the study. The students' parents signed the consent forms (whether participating or refusing participation in the study). The researcher informed them about the purpose of the study, instructional process, and ethical principles that directed the study. The teachers and school principals were informed that the preliminary and detailed conclusions of the study will be shared with them. The first researcher conducted the lessons in all groups because when the study was conducted, the use of augmented reality technology in education was very new. Also, the science teachers had inadequate knowledge and experience about augmented reality applications. Before the study, the researcher attended the classes with the teacher and made observations to help students become accustomed to the presence of the researcher in the classroom. This prolonged engagement increased the reliability of the implementation of instructions and data collection (Guba & Lincoln, 1989).

3.2 Instructional Activities

The 3-week, 19-h intervention process, which was based on the prepared lesson plan, was conducted. The students participated in the learning process that included both augmented reality and argumentation activities about astronomy.

3.2.1 Augmented Reality Activities

Free applications such as Star Chart, Sky View Free, Aurasma, Junaio, Augment, and i Solar System were used with students' tablet computers in augmented reality instructions. Tablet computers were provided by the school administration from their stock. Videos, simulations, and 3D visuals generated by applications were used as "overlays." In addition, pictures, photographs, colored areas in the worksheets, and students' textbooks were used as "trigger images." Students had the opportunity to interact with and manipulate these videos, simulations, and 3D visuals while using the applications. With applications such as Sky View Free and Star Chart, students were provided with the resources to make sky observations.

A detailed description of the activities used in augmented reality is given in Appendix Table 8.

3.2.2 Argumentation Activities

Before the instruction, the students were divided into six groups by the teacher, paying attention to heterogeneity in terms of gender and academic achievement. After small group discussions, the students participated in whole-class discussions. Competing theories cartoons, tables of statements, constructing an argument, and argument-driven inquiry (ADI) frameworks were used to support argumentation in the learning process. Argument-driven inquiry consists of eight steps including the following: identification of the task, the generation and analysis of data, the production of a tentative argument, an argumentation session, an investigation report, a double-blind peer review, revision of the report, and explicit and reflective discussion (Sampson & Gleim, 2009; Sampson et al., 2011).

A detailed description of the activities used in argumentation is given in Appendix Table 9.

4 Data Collection

The data were collected through unstructured and participant observations (Maykut & Morehouse, 1994; Patton, 2002). The instructional intervention was recorded with a video camera, and the students' argumentation processes were also recorded with a voice recorder.

Since all students spoke at the same time during group discussions, the observation records were insufficient to understand the student talks. To determine what each student in the group said during the argumentation process, a voice recorder was placed in the middle of the group table, and a voice recording was taken throughout the lesson. A total of 2653.99 min of voice recordings were taken in the six groups.

4.1 Data Analysis

The analysis of the data was conducted with inductive and deductive approaches. Before coding, the data were arranged. The critical thinking data were organized by day. The argumentation skills were organized by day and also on the basis of the groups. After generating codes during the inductive analysis of the development of critical thinking skills, a deductive approach was adopted (Patton, 2002). The critical thinking skills dimensions discussed by Ennis (2011) and Ennis (1991) were used to determine the relationship between codes. Ennis (2011) prepared an outline to distinguish critical thinking dispositions and skills by synthesizing of many years of studies. These critical skills that contain abilities that ideal critical thinkers have were used to generate codes from students' talks. This skills and abilities were given in Appendix Table 10. Then "clarification skills, decision making-supporting skills, inference skills, advanced clarification skills, and other/strategy and techniques skills" discussed by Ennis (1991) and Ennis (2011) were used to determine the categories. The change in the argumentation abilities of the students was analyzed descriptively based on the Toulmin argument model (Toulmin, 1990) using the data obtained from the students' voice recordings. The argument structures of each group during verbal argumentation were determined by dividing them into components according to the Toulmin model (Toulmin, 1990). The first three items (data, claim, and warrant) in the Toulmin model form the basis of an argument, and the other three items (rebuttal, backing, and qualifier) are subsidiary elements of the argument (Toulmin, 1990).

Some quotations regarding the analysis of the arguments according to the items are given in Appendix Table 11.

Arguments from the whole group were put into stages based on the argumentation-level model developed by Erduran et al. (2004) to examine the changes in each lesson and to make comparisons between the small groups of students. By considering the argument model developed by Toulmin, Erduran et al. (2004) created a five-level framework for the assessment of the quality of argumentation supposing that the quality of the arguments including rebuttals was high. The framework is given in Table 2.

4.2 Validity and Reliability

To confirm the accuracy and validity of the analysis, method triangulation, triangulation of data sources, and analyst triangulation were used (Patton, 2002).

Table 2 The framework for the assessment of the quality of argumentation (Erduran et al., 2004; pp. 928)

Levels	Description
Level 1	Level 1 argumentation consists of arguments that are a simple claim versus a counterclaim or a claim versus claim
Level 2	Level 2 argumentation has arguments consisting of claims with either data, warrants, or backings, but do not contain any rebuttals
Level 3	Level 3 argumentation has arguments with a series of claims or counterclaims with either data, warrants, or backings with the occasional weak rebuttal
Level 4	Level 4 argumentation shows arguments with a claim with a clearly identifiable rebuttal. Such an argument may have several claims and counterclaims as well, but this is not necessary
Level 5	Level 5 argumentation displays an extended argument with more than one rebuttal

For analyst triangulation, the qualitative findings were also analyzed independently by a researcher studying in the field of critical thinking and argumentation, and then these evaluations made by the researchers were compared.

Video and audio recordings of intervention and documents from the activities were used for the triangulation of data sources. In addition, the data were described in detail without interpretation. Additionally, within the reliability and validity efforts, direct quotations were given in the findings. In this sense, for students, codes such as S1, S2, and S3 were used, and the source of data, group number, and relevant date of the conversation were included at the end of the quotations.

In addition, experts studying in the field of critical thinking and argumentation were asked to verify all data and findings. After the process of reflection and discussion with experts, the codes, subcategories, and categories were revised.

For reliability, some of the data randomly selected from the written transcripts of the students' audio recordings were also coded by a second encoder, and the interrater agreement between the two coders, determined by Cohen's kappa (Cohen, 1960), was $\kappa = 0.86$, which is considered high reliability.

5 Results

5.1 Development of Critical Thinking Ability

The development of critical thinking skills was given separately for the trend drastically changed on the day when the first skills were used by the students. All six dimensions of critical thinking skills were included in students' dialogs or when there was a decrease in the number of categories of critical thinking skills.

The codes, subcategories, and categories of critical thinking skills that occurred on the first day (dated 11.05) are given in Table 3.

Clarification skills, inference skills, other/strategy and technical skills, advanced clarification skills, and decision-making/supporting skills occurred on the first day. The students mostly used decision-making/supporting skills ($f = 55$). Under the decision-making/supporting skills category, students mostly explained observation data ($f = 37$). S7, S1, and S20 stated the data they presented about their observations with the Star Chart and Sky View applications as follows:

S7: Venus is such a yellowish reddish colour.

S1: What was the colour? Red and big. The moon's color is white.

S20: Not white here.

S1: How?

S20: It's not white here. (Audio Recordings (AuR), Group 2 / 11.05).

Additionally, S19 mentioned the observation data with the words "I searched Saturn. It is bright. It does not vibrate. It is yellow and it's large." (AuR, Group 2 / 11.05).

Decision-making/supporting skills were followed by inference ($f = 17$), clarification ($f = 13$), advanced clarification ($f = 5$), and skills and other/strategy technical skills ($f = 1$).

In Table 4, the categories, subcategories, and codes for critical thinking skills that occurred on the fifth day (dated 18.05) are presented.

It was observed for the first time on the fifth day that *all six dimensions of critical thinking skills* were included in students' dialogs. These are, according to the frequency of use, inference ($f = 152$), decision-making/support ($f = 116$), clarification

Table 3 The codes, subcategories, and categories of critical thinking skills that occurred on the first day

Categories	Subcategories	*f* (frequency)
Decision making-supporting skills	Explaining observation data	37
	Giving reasons	11
	Judging observation data	3
	Seeking precision	2
	Judging the credibility	1
	Using credible sources	1
Inference skills	Making inference from the available data	6
	Making counter-claim	5
	Making claim	5
	Using evidence to support the claim	1
	Making alternative explanations inconsistent with facts	1
Clarification skills	Asking questions of clarification the situation	5
	Asking for clarification	4
	Asking for example	1
	Asking for comparison	1
	Asking for reason	1
	Summarizing	1
Advanced clarification skills	Making comparison	4
	Giving example	1
Other/strategy and technique skills	Giving solutions to problems	1

($f=43$), advanced clarification ($f=8$), other/strategy and technique ($f=3$), and suppositional thinking and integrational ($f=2$) skills.

On this date, *judging the credibility of the source* from decision-making/supporting skills ($f=1$) was the skill used *for the first time.*

Unlike other days, for the first time, a student tried to prove his thoughts *with an analogy* in advanced clarification skills. An exemplary dialogue to this finding is as follows:

S19: Even the Moon remains constant, we will see different faces of the moon because the Earth revolves around its axis.

S6: I also say that it turns at the same speed. So, for example, when this house turns like this while we return in the same way, we always see the same face. (AuR, 18.05, Group 2).

Here, S6 tried to explain to his friend that they always see the same face of the moon by comparing how they see the same face of the house.

In Table 5, the categories, subcategories, and codes for critical thinking skills that occurred on the sixth day (dated 21.05) are included.

There is *a decrease* in the number of categories of critical thinking skills. It was determined that the students used mostly inference skills in three categories ($f=38$). Additionally, students used decision-making/support ($f=34$) and clarification ($f=9$) skills. In inference skills, it is seen that students often make claims ($f=33$) and rarely infer from the available data ($f=4$).

Among the decision-making/support skills, students mostly used the skill to give reasons ($f=28$). S24 accepted herself as Uranus during the activity, and she gave reason to make Saturn as an enemy like that: "No, Saturn would be my enemy too. Its ring is more

Table 4 The categories, subcategories, and codes for critical thinking skills that occurred on the fifth day

Categories	Subcategories	Codes	*f*
Inference skills	Making claim		98
	Making counter-claim		27
	Making prediction		15
	Using inductive reasoning		6
	Using deductive reasoning		2
	Changing first claim		2
	Making alternative explanations inconsistent with facts		2
Decision making-supporting skills	Giving reasons	Giving reason for the claim	61
		Using evidence for the claim	7
		Giving reason for disagreements	3
		Giving reason for the question asked	3
	Explaining observation data		34
	Judging the accuracy of the statement		6
	Judging the credibility		1
	Using credible sources		1
Clarification skills	Asking friend about his/her opinion		26
	Asking for reason		9
	Asking detailed explanation		5
	Trying to understand the explanation		2
	Asking questions of clarification the situation		1
Advanced clarification skills	Making comparison		5
	Giving example		2
	Trying to prove with analogy		1
Suppositional thinking and integration skills	Thinking from a different perspective		2
Other/strategy and technique skills	Be sensitive to the ideas of others		2
	Giving solutions to problems		1

Table 5 The categories, subcategories, and codes for critical thinking skills that occurred on the sixth day

Categories	Subcategories	Codes	*f*
Inference skills	Making claim		33
	Making inference from the available data		4
	Rejecting the judgment		1
Decision making-supporting skills	Giving reasons	Giving reason for the claim	26
		Using evidence for the claim	1
		Giving reason for disagreements	1
	Judging the accuracy of the statement		3
	Explaining observation data		2
	Using credible sources		1
Clarification skills	Asking friend about his/her opinion		4
	Asking questions of clarification the situation		3
	Asking for reason		2

distinctive, it can be seen from the Earth, its ring is more beautiful than me." (AuR, 21.05, Group 3/).

The categories, subcategories, and codes for critical thinking skills that occurred on the ninth day (dated 28.05) are presented in Table 6.

In the course of the day dated 28.05, *six categories* of critical thinking skills were observed: clarification, inference, other/strategy and technique, advanced clarification, decision-making/support, suppositional thinking and integration skills. Furthermore, the subcategories under these categories are also *very diverse.*

There are 10 subcategories under clarification skills ($f=57$), which are the most commonly used skills. The frequency of using these skills is as follows: asking his friend about his opinion ($f=15$), asking questions to clarify the situation ($f=12$), explaining his statement ($f=10$), summarizing the solutions of other groups ($f=7$), asking for a detailed explanation ($f=4$), summarizing the idea ($f=3$), explaining the solution proposal ($f=2$), asking for a reason ($f=2$), focusing on the question ($f=1$), and asking what the tools used in experiment do ($f=1$) skills. Explaining the solution proposal, asking what the tools used in the experiment do, and focusing on the question are *the first skills* used by the students.

When the qualitative findings regarding the critical thinking skills of the students were examined as a whole, it was determined that there was an improvement in the students' critical thinking skills dimensions in the lessons held in the first 5 days (between 11.05 and 18.05). There was a decrease in the number of critical thinking skills dimensions in the middle of the intervention (21.05). However, after this date, there was an increase again in the number of critical thinking skills dimensions; and on the last day of the intervention, all the critical thinking skills dimensions were used by the students. In addition, it was determined that the skills found under these dimensions showed great variety at this date. Only in the middle (18.05) and on the last day (28.05) of the intervention did students use the skills in the six dimensions of critical thinking.

Table 6 The categories, subcategories, and codes for critical thinking skills that occurred on the ninth day

Categories	Subcategories	Codes	*f*
Clarification skills	Asking friend about his/her opinion		15
	Asking questions of clarification the situation		12
	Explaining the statement		10
	Summarizing the solutions of other groups		7
	Asking detailed explanation		4
	Summarizing the idea		3
	Explaining the solution proposal		2
	Asking for reason		2
	Focusing on the question		1
	Asking what the tools used in experiment do		1
Inference skills	Making counter-claim		16
	Making prediction		13
	Using deductive reasoning		7
	Making claim		2
	Rejecting the judgment		1
Other/strategy and technique skills		Giving solutions to problems	34
		Trying to make a common decision	5
Decision making-supporting skills	Giving reason	Giving reason for disagreements	20
		Giving reason for the claim	8
		Explaining observation data	2
		Giving reason for possible counter claims	1
		Giving reasons for the question asked	1
	Judging the accuracy of the statement		1
Suppositional thinking and integration skills	Considering and reasoning from other disagreed propositions		6
	Thinking from a different perspective		1
Advanced clarification skills	Giving example		4
	Explaining differences		1

It was determined that students used *mostly decision-making/support, inference, and clarification skills.* According to the days, it was determined that the students mostly used *inference skills* (12.05, 15.05, 18.05, and 21.05) among these skills.

5.2 The Argumentation Abilities of the Students

5.2.1 Argument Structures in Students' Verbal Argumentation Activities

Instead of the argument structures of all groups, only an example of one group is presented because of including both basic and subsidiary items in the Toulmin argument model. In Table 7, the argument structures in the verbal argumentation activities of the fourth group of students are presented due to the use of the "rebuttal" item.

When the argument structures in the verbal argumentation process of the six groups were examined, it was found that all groups engaged in the argumentation and produced arguments. In the activities, students mostly made claims. This was followed by data and warrant items. In the "the phases of the moon" activity, it was determined that only the second and fourth groups used rebuttal and the other groups did not.

The number of rebuttals used by the groups is lower in "the planets-table of statements" activity than in other activities. The rebuttals used are also weak. The use of rebuttals differs in the "who is right?" and "urgent solution to space pollution" activities. The number of rebuttal students used in these activities is higher than that in the other activities. The quality rebuttals are also higher.

When the structure of the warrants is examined, there are more unscientific warrants in the "urgent solution to space pollution" and "who is right" activities, while the correct scientific and partially correct scientific warrants were more frequently used in the "the phases of the moon" and "the planets table of statements" activities.

When the models related to the argument structures are examined in general, it was found that there is a decrease in the type of items used and the number of uses in the "the phases of the moon" and "the planets-table of statements" activities rather than the "urgent solution to space pollution" and "who is right" activities.

When the results were analyzed in terms of groups, it was determined that the argument structures of the second and fourth groups showed more variety than those of the other groups.

5.2.2 The Change of Argumentation Levels

The argumentation levels achieved by six groups created in the "who is right," " the planets-table of statements," "phases of the moon," and "urgent solution to space pollution" activities are shown in Fig. 2.

In the first verbal argumentation activity, "who is right?," the arguments achieved by the five of the six groups were at level 5. Additionally, the arguments achieved by one group, which was group 6, were at level 4.

In the second verbal argumentation activity "table of statements," a decrease was determined at the levels of the argumentation of the other groups except group 1 and group 3. In the "the phases of the moon" activity, there was a decrease at the level of argumentation achieved by the other groups except for group 2 and group 4. In the last argumentation activity, "urgent solution to space pollution," it was found that the arguments of all groups were at level 5.

Table 7 The argument structures in the verbal argumentation activities of the fourth group of students

The verbal argumentation activities	The items in the Toulmin argument model	The subitems	*f*
Who is right?	Counter-claim		6
	Claim		4
	Data		9
	Rebuttal	Weak rebuttal	4
		Qualified rebuttal	2
	Warrant	Unscientific warrant	5
		Partially correct scientific warrant	1
	Qualifier		2
Table of statements	Claim		10
	Data		8
	Warrant	Scientific warrant	2
		Incorrect inference	2
		Partially correct scientific warrant	1
	Qualifier		2
	Rebuttal	Qualified rebuttal	1
The phases of the moon	Claim		18
	Counter-claim		4
	Warrant	Incorrect inference	4
		Scientific warrant	1
		Partially correct scientific warrant	1
	Rebuttal	Qualified rebuttal	2
		Incorrect rebuttal	1
Urgent solution to space pollution	Claim		16
	Counter-claim		1
	Rebuttal	Weak rebuttal	6
		Qualified rebuttal	6
	Warrant	Scientific warrant	2
		Unscientific warrant	2
		Partially correct warrant	2

6 Conclusions and Discussion

The critical thinking skills of the students developed until the middle of the intervention, and the frequency of using critical thinking skills varied after the middle of the intervention. When the activities in the lessons were examined, on the days when critical thinking skills were frequently used, activities including argumentation methods were performed. Based on this situation, it could be revealed that the frequency of using critical thinking skills by students varies according to the use of the argumentation method.

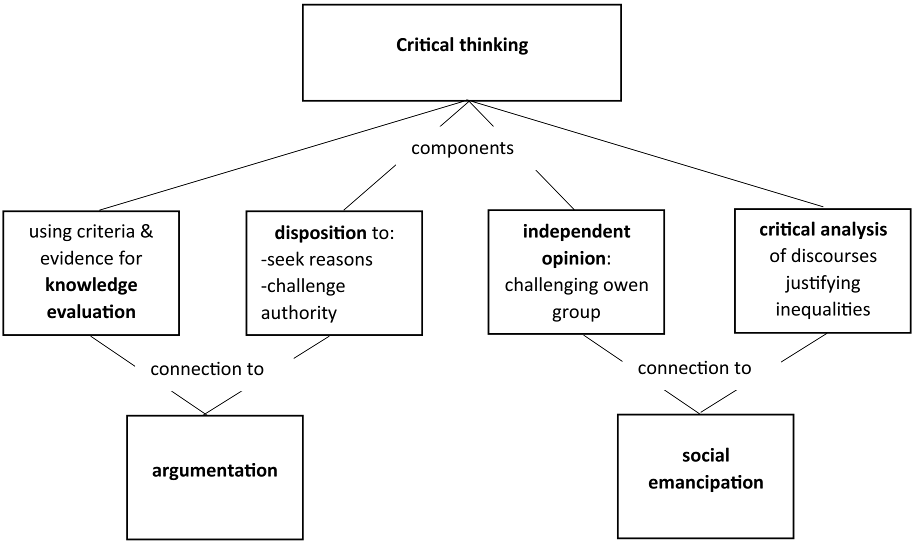

Fig. 2 A characterization of the components of critical thinking (Jiménez-Aleixandre & Puig, 2012, p. 6)

Argumentation is defined as the process of making claims about a scientific subject, supporting them with data, providing reasons for proof, and criticizing, rebutting, and evaluating an idea (Toulmin, 1990). According to the definition of argumentation, these processes are also in the subdimensions of critical thinking skills. The ability to provide reasons for critical thinking skills in decision-making/supporting skills is the equivalent of providing reasons for proof in the argumentation process using warrants in the Toulmin argument model. Different types of claims under inference skills are related to making claims in the argumentation process, and rejecting a judgment is related to rebutting an idea in the argumentation process. In this context, the argumentation method is thought to contribute to the development of critical thinking skills within AR.

Another qualitative finding reached in the study is that the skills most used in the subdimensions differ according to the days. This can be explained by the different types of activities performed in each lesson. For example, on the day when the ability to explain observation data was used the most, students observed the sky, constellations, and galaxies with the Star Chart or Sky View applications or observed the planets with the i-Solar System application, and they presented the data they obtained during these observations.

Regarding the verbal argumentation structure of the groups, the findings imply that all groups engaged in argumentation and produced arguments. This finding presented evidence with qualitative data to further verify Squire & Jan's (2007) research conducted with primary, middle, and high school students to investigate the potential of a location-based AR game in environmental science concluding that all groups engaged in argumentation. Similarly, Jan (2009) investigated the experience of three middle school students and their argumentative discourse on environmental education using a location-based AR game, and it was found that all students participated in argumentation and produced arguments.

Another finding in the current study was that students mostly made claims in the activities. This situation can be interpreted as students being strong in expressing their opinions. Similar findings are found in the literature (Author, 20xxa; Cavagnetto et al., 2010; Erduran et al., 2004; Novak & Treagust, 2017). In addition, it was concluded that the students failed to use warrants and data, they could not support their claims with the data, and they did not use "rebuttal" in these studies. However, in this study in which both augmented reality applications and argumentation methods were used, students mostly made contradictory claims and used data and warrants in their arguments. This situation can be interpreted as students being strong in defending their opinions. Additionally, although it was stated in many of the studies that students' argumentation levels were generally at level 1 or level 2 (Erdogan et al., 2017; Erduran et al., 2004; Venville & Dawson, 2010; Zohar & Nemet, 2002), it was found that most of the students' arguments were at level 4 and level 5 in the current study. Arguments are considered to be high quality in line with the existence of rebuttals, and discussions involving rebuttals are characterized as having a high level of argumentation (Aufschnaiter et al., 2008; Erduran et al., 2004). Students used rebuttals in their arguments, and their arguments were at high levels, which indicates that students could produce quality arguments. The reason for these findings to differ from those of other studies may be due to the augmented reality technology used in the current study. Enriched representations make it easier to see the structure of arguments (Akpınar et al., 2014), helping students to improve their awareness, increase the number of words they use and comments they make (Erkens & Janssen, 2006), and provide important information about the subject (Clark et al., 2007). By observing enriched representations, students collect evidence for argumentation (Clark & Sampson, 2008) and explore different points of view to support their claim (Oestermeier & Hesse, 2000). AR technology, which includes enriched representations, may have increased the accessibility of rich data to support students' arguments; and using these data has helped them to support their arguments and enabled them to discover different perspectives. For example, S4 explained that the statement in the table is incorrect because she observed Uranus, Jupiter, and Neptune having rings around them in the application "I-solar system" as Uranus. She used the data obtained in the AR application to support her claim.

When the models related to the argument structures are examined in general, it was concluded that the type of items, the number of items, and the rebuttals used in scientific activities were less than those in the activities involving socioscientific issues. The rebuttals used were also weak. There are also findings in the literature that producing arguments on scientific issues is more difficult than producing arguments on socioscientific issues (Osborne et al., 2004).

When the structure of the warrants in the students' arguments was examined, it was seen that there are more nonscientific warrants in socioscientific activities, and the scientific and partially scientific warrants are more in the activities that contain scientific subjects. This shows that students were unable to combine what they have learned in science with socioscientific issues. Albe (2008) and Kolsto (2001) stated that scientific knowledge is very low in students' arguments on socioscientific issues. Similarly, the results of the studies conducted in the related literature support this view (Demircioglu & Ucar, 2014; Sadler & Donnelly, 2006; Wu & Tsai, 2007).

When the argument structures in the activities are analyzed by groups, the argument structures of the two groups vary more than the other groups, and the argumentation levels of these groups are at level 4 and level 5. This might be because some students have different prior knowledge about subjects. Different studies have also indicated that content knowledge plays an important role in the quality of students' arguments (Acar, 2008; Aufschnaiter et al., 2008; Clark & Sampson, 2008; Cross et al., 2008; Sampson & Clark, 2011). In many studies, it has been emphasized that the most important thing affecting the choice and process of knowledge is previous information (Stark et al., 2009). To better understand how previous information affects argumentation quality in astronomy education, investigating the relationship between middle school students' content knowledge and argumentation quality could be a direction of future research.

7 Limitations and Future Research

There are some limitations in this study. First, this study was implemented in a private school. Therefore, the results are true for these students. Future research is necessary to be performed with the students in public schools. Second, the researcher conducted the lessons because the science teacher had no ability to design AR learning practices. Teachers and students creating their own AR experiences is an important way to bring the learning outcomes of AR available to a wider audience (Romano et al., 2020). Further research can be conducted in which the science teacher of the class is the instructor. Another limitation of the study is that the instruction with AR-based argumentation was time-consuming, and the time allocated for the "Solar System and Beyond" unit in the curriculum was not sufficient for the implementation, because students tried to understand to use AR applications, and they needed time to reflect on the activities despite prior training on AR before the instructional process. This situation may cause cognitive overload (Alalwan et al., 2020). The adoption and implementation of educational technologies are more difficult and time-consuming than other methods (Parker & Heywood, 1998). A longer period is needed to prepare student-centered and technology-supported activities.

Appendix

Tables 8, 9, 10 and 11

Table 8 The activities performed with augmented reality technology

Activities	Content	AR applications used in activities
My constellation story	Designing a constellation, preparing a poster with information about this constellation, creating a story about the constellation, recording the narration of this story with video and superimposing the video on the poster through Aurasma	Aurasma
Meteor shower	Watching a video of a meteor shower superimposed on textbook	Aurasma
The moon and planets	Observing three-dimensional images of the moon and planets superimposed on a textbook	Blender and Aurasma
Space shuttle and the moment the shuttle launches	Observing a 3D image of the space shuttle with the Augment app. and the first launch moment of the shuttle superimposed on a textbook with Aurasma	Augment and Aurasma
Moon, Earth, telescope, space shuttle	Observing the rotation of the moon in its orbit around the Earth, the 3D telescope, and the space shuttle view	Augment
The planets	Exploring 3D models, videos, images, and sounds about planets in the "Augmented Reality Magic Book" created by Nedim Slijepcevic and Wanju Huang	Junaio
Solar system	Interactively observing the solar system	i Solar System book and its application
First landing on the moon	Examining the first landing on the moon while this is happening in front of you in an immersive virtual world	Moon walking
Sky observation	Observing the sky (the current position of every star and planet visible from the Earth and where they are and 3D effects, distances, brightness, and positions of stars, constellations, and planets)	Star Chart, Sky View

Table 9 Activities performed with argumentation

Activities	Content	Argumentation frameworks
Who is right?	To engage in argumentation on the question of whether astrology is a science or not Students were presented with two competing theories in the form of a cartoon. They were asked to indicate the one they believe in and argue why they thought they were correct	Competing theories-cartoons
The planets-table of statements	To engage in argumentation whether the statements in the presented table about the planets are true or false Students were given a table with statements about planets. They were asked to indicate whether these statements were correct or incorrect and to express their opinions with data and warrants	Table of statements
The phases of the moon	To explain the following: What are the phases of the moon and why do we see them in the order we do? Why do we see the same side of the moon every day?	Argument-driven inquiry (ADI) The ADI steps were explained in "3.2.2 Argumentation activities" section
Urgent solution to space pollution	Making arguments about preventing space pollution Students were given an explanation of space pollution and a case about space pollution. Then, they discussed about the solutions to space pollution and which data statements provide the strongest explanation for the phenomenon	Constructing an argument

Table 10 The critical thinking skills and abilities (Ennis, 2011, pp. 2–4)

Critical thinking skills	Abilities	
Basic clarification	1. Focus on a question	a. Identify or formulate a question b. Identify or formulate criteria for judging possible answers c. Keep the question and situation in mind
	2. Analyze arguments	a. Identify conclusions b. Identify reasons or premises c. Ascribe or identify simple assumptions (see also ability 10) c. Identify and handle irrelevance d. See the structure of an argument e. Summarize
	3. Ask and answer clarification and/or challenge questions, such as	a. Why? b. What is your main point? c. What do you mean by·? d. What would be an example? e. What would not be an example (though close to being one)? f. How does that apply to this case(describe a case, which appears to be a counterexample)? g. What difference does it make? h. What are the facts? i. Is this what you are saying:__________________? j. Would you say more about that?
Two bases for a decision	4. Judge the credibility of a source Major criteria (but not necessary conditions)	a. Expertise b. Lack of conflict of interest c. Agreement with other sources d. Reputation e. Use of established procedures f. Known risk to reputation (the source's knowing of a risk to reputation, if wrong) g. Ability to give reasons h. Careful habits

Table 10 (continued)

Critical thinking skills	Abilities	
	5. Observe and judge observation reports. Major criteria (but not necessary conditions, except for the first)	a. Minimal inferring involved b. Short time interval between observation and report c. Report by the observer, rather than someone else (that is, the report is not hearsay) d. Provision of records e. Corroboration f. Possibility of corroboration g. Good access h. Competent employment of technology, if technology applies i. Satisfaction by observer (and reporter, if a different person) of the credibility criteria in Ability# 4 above
Inference	6. Deduce and judge deduction	a. Class logic b. Conditional logic c. Interpretation of logical terminology, including (1) Negation and double negation (2) Necessary and sufficient condition language (3) Such words as "only," "if and only if," "or," "some," "unless," and "not both" d. Qualified deductive reasoning(a loosening for practical purposes)

Table 10 (continued)

Critical thinking skills	Abilities	
	7. Make material inferences (roughly "induction")	a. To generalizations. Broad considerations: (1) Typicality of data, including valid sampling where appropriate (2) Volume of instances (3) Conformity of instances to generalization (4) Having a principled way of dealing with outliers b. To explanatory hypotheses(IBE: "inference-to-best explanation"): (1) Major types of explanatory conclusions and hypotheses: (a) Specific and general causal claims (b) Claims about the beliefs and attitudes of people (c) Interpretation of authors' intended meanings (d) Historical claims that certain things happened (including criminal accusations) (e) Reported definitions (f) Claims that some proposition is an unstated, but used, reason (2) Characteristic investigative activities (a) Designing experiments, including planning to control variables (b) Seeking evidence and counterevidence, including statistical significance (c) Seeking other possible explanations (3) Criteria, the first four being essential, the fifth being desirable (a) The proposed conclusion would explain or help explain the evidence (b) The proposed conclusion is consistent with all known facts (c) Competitive alternative explanations are inconsistent with facts (d) A competent sincere effort has been made to find supporting and opposing data and alternative hypotheses (e) The proposed conclusion seems plausible and simple, fitting into the broader picture
	8. Make and judge value judgments Important factors	a. Background facts b. Consequences of accepting or rejecting the judgment c. Prima facie application of acceptable principles d. Alternatives e. Balancing, weighing, deciding

Table 10 (continued)

Critical thinking skills	Abilities	
Advanced clarification	9. Define terms and judge definitions, using appropriate criteria	a. Definition form. (1) Synonym (2) Classification (3) Range (4) Equivalent-expression (5) Operational (6) Example and non-example b. Definitional functions (acts) (1) Report a meaning (criteria: the five for an explanatory hypothesis) (2) Stipulate a meaning (criteria: convenience, consistency, avoidance of impact equivocation) (3) Express a position on an issue(positional definitions, including "programmatic" and "persuasive" definitions) Criteria: those for a position c. Content of the definition d. Identifying and handling equivocation
Supposition and integration	10.Consider and reason from premises, reasons, assumptions, positions, and other propositions with which they disagree or about which they are in doubt, without letting the disagreement or doubt interfere with their thinking	
	11. Integrate the dispositions and other abilities in making and defending a decision	
Auxiliary abilities	12. Proceed in an orderly manner appropriate to the situation	a. Follow problem-solving steps b. Monitor their own thinking (that is, engage in metacognition) c. Employ a reasonable critical thinking checklist
	13.Be sensitive to the feelings, level of knowledge and degree of sophistication of others	
	14. Employ appropriate rhetorical strategies in discussion and presentation (oral and written),including employing and reacting to "fallacy" labels in an appropriate manner. Examples of fallacy labels are "circularity," "bandwagon," "post hoc," "equivocation," "non sequitur," and "straw person"	

Table 11 Quotations regarding the analysis of the arguments according to the items

Items	Subdimensions	Quotations
Claim		“Astrology is not science” (AuR (Audio Recordings),12.05, Group 1, S25 /00.00–05.32)
		“The first planet to be encountered when leaving Earth is Venus, the last planet Mercury. False.”(AuR,18.05, Group 3, S21/14.33–17.24)
Counterclaim		“I do not think so” (AuR, 12.05, Group 4, S16 /00.00–04.59)
		“No, that’s right of course” (AuR, 18.05, Group 6, S15 /14.29–20.06)
Data		“At the same time, 2000 people born on the same day and at the same time were examined in a research conducted in the past. But there is no similarity between them.” (AuR, 12.05, Group 1, S4 /00.00–05.32)
Warrant	Scientific warrant	“Because I think Mars is the first planet while travelling from Earth to this side, and the last planet is Neptune.” (AuR, 18.05, Group 1, S4 /14.49–21.02)
	Unscientific warrant	“They did scientific research and they concluded that it is true. What do you say Ö8? It may be wrong scientifically. Because perhaps someone who was hostile to astrology bribed the man who did this research (Non-scientific justification). How do you know?” (AuR, 12.05, Group 4, S6-S8 /00.00–04.59)
	Incorrect inference	“I think the phases of the moon are due to the amount of light reflected by the Sun.” (AuR, 22.05, Group 5, S23 /16.23–19.25)
Qualifier		“I do not think so. Sometimes, it can be so different.” (AuR, 12.05., Group 4, S16/00.00–04.59)
Rebuttal	Qualified rebuttal	“Dear Friends, what you say is absolutely wrong because we always see the same face of the moon as the Earth rotates.”
	Weak rebuttal	“Neptune is not the farthest. The answer to the first question is wrong because I think Neptune is not the farthest planet. Mars is the closest to the Earth.”(AuR, 18.05., Group 3, S13-S14/14.33–17.24)
	Incorrect rebuttal	“There is a tiny time difference in the rotational speed of the Earth and the Moon. There is a slight deviation in the rotational speed of the Earth and Moon per hour. That’s why we don’t see the same face of the moon.” (AuR, 18.05., Group 3, S14/19.09–20.27)

Funding This study is a part of Tuba Demircioğlu's dissertation supported by the Cukurova University Scientific Research Projects (grant number: SDK20153929).

The manuscript is part of first author's PhD dissertation. The study was reviewed and approved by the PhD committee in the Cukurova University Faculty of Education, as well as by the committee of Ministry of National Education. The parents of students were provided with written informed consent.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

Acar, O. (2008). *Argumentation skills and conceptual knowledge of undergraduate students in a physics by inquiry class*. (Doctoral dissertation, The Ohio State University). https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu1228972473&disposition=inline

Akpınar, Y., Ardaç, D., & Er-Amuce, N. (2014). Development and validation of an argumentation based multimedia science learning environment: Preliminary findings. *Procedia-Social and Behavioral Sciences, 116*, 3848–3853.

Aktamış, H., & Arıcı, V. A. (2013). The effects of using virtual reality software in teaching astronomy subjects on academic achievement and retention. *Mersin University Journal of the Faculty of Education, 9*(2), 58–70.

Alalwan, N., Cheng, L., Al-Samarraie, H., Yousef, R., Alzahrani, A. I., & Sarsam, S. M. (2020). Challenges and prospects of virtual reality and augmented reality utilization among primary school teachers: A developing country perspective. *Studies in Educational Evaluation, 66*, 100876.

Albe, V. (2008). When scientific knowledge, daily life experience, epistemological and social considerations intersect: Students' argumentation in group discussions on a socio-scientific issue. *Research in Science Education, 38*(1), 67–90.

Arici, F., Yildirim, P., Caliklar, Ş., & Yilmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education, 142*, 103647.

Atwood-Blaine, D., & Huffman, D. (2017). Mobile gaming and student interactions in a science center: The future of gaming in science education. *International Journal of Science and Mathematics Education, 15*(1), 45–65.

Baran, B., Yecan, E., Kaptan, B., & Paşayiğit, O. (2020). Using augmented reality to teach fifth grade students about electrical circuits. *Education and Information Technologies, 25*(2), 1371–1385.

Billinghurst, M. (2002). Augmented reality in education. *New Horizons for Learning, 12*(5), 1–5.

Bressler, D. M., & Bodzin, A. M. (2013). A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. *Journal of Computer Assisted Learning, 29*(6), 505–517.

Buckingham Shum, S. J., MacLean, A., Bellotti, V. M., & Hammond, N. V. (1997). Graphical argumentation and design cognition. *Human-Computer Interaction, 12*(3), 267–300.

Cadmus, R. R., Jr. (1990). A video technique to facilitate the visualization of physical phenomena. *American Journal of Physics, 58*(4), 397–399.

Cai, S., Chiang, F. K., & Wang, X. (2013). Using the augmented reality 3D technique for a convex imaging experiment in a physics course. *International Journal of Engineering Education, 29*(4), 856–865.

Chai, C. S., Deng, F., Tsai, P. S., Koh, J. H. L., & Tsai, C. C. (2015). Assessing multidimensional students' perceptions of twenty-first-century learning practices. *Asia Pacific Education Review, 16*(3), 389–398.

Cai, S., Liu, C., Wang, T., Liu, E., & Liang, J. C. (2021). Effects of learning physics using augmented reality on students' self-efficacy and conceptions of learning. *British Journal of Educational Technology, 52*(1), 235–251.

Cavagnetto, A., Hand, B. M., & Norton-Meier, L. (2010). The nature of elementary student science discourse in the context of the science writing heuristic approach. *International Journal of Science Education, 32*(4), 427–449.

Chang, H. Y., Wu, H. K., & Hsu, Y. S. (2013). Integrating a mobile augmented reality activity to contextualize student learning of a socioscientific issue. *British Journal of Educational Technology, 44*(3), 95–99.

Chang, H. Y., Liang, J. C., & Tsai, C. C. (2020). Students' context-specific epistemic justifications, prior knowledge, engagement, and socioscientific reasoning in a mobile augmented reality learning environment. *Journal of Science Education and Technology, 29*(3), 399–408.

Chen, S. Y., & Liu, S. Y. (2020). Using augmented reality to experiment with elements in a chemistry course. *Computers in Human Behavior, 111*(2020), 106418.

Chen, C. H., Yang, J. C., Shen, S., & Jeng, M. C. (2007). A desktop virtual reality earth motion system in astronomy education. *Educational Technology & Society, 10*(3), 289–304.

Chen, C. H., Huang, C. Y., & Chou, Y. Y. (2019). Effects of augmented reality-based multidimensional concept maps on students' learning achievement, motivation and acceptance. *Universal Access in the Information Society, 18*(2), 257–268.

Chen, C. C., Chen, H. R., & Wang, T. Y. (2022). Creative situated augmented reality learning for astronomy curricula. *Educational Technology & Society, 25*(2), 148–162.

Cheng, K. H., & Tsai, C. C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology, 22*(4), 449–462.

Chiang, T. H. C., Yang, S. J. H., & Hwang, G. J. (2014). An augmented reality-based mobile learning system to improve students' learning achievements and motivations in natural science inquiry activities. *Journal of Educational Technology & Society, 17*(4), 352–365.

Clark, D. B., & Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching, 45*(3), 293–321.

Clark, D. B., Stegmann, K., Weinberger, A., Menekse, M., & Erkens, G. (2007). Technology-enhanced learning environments to support students' argumentation. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 217–243). Springer.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Cross, D., Taasoobshirazi, G., Hendricks, S., & Hickey, D. T. (2008). Argumentation: A strategy for improving achievement and revealing scientific identities. *International Journal of Science Education, 30*(6), 837–861.

Demircioglu, T., & Ucar, S. (2014). Investigation of written arguments about Akkuyu nuclear power plant. *Elementary Education Online, 13*(4), 1373–1386.

Demircioglu, T., & Ucar, S. (2015). Investigation the effect of argument-driven inquiry in laboratory instruction. *Educational Sciences: Theory and Practice, 15*(1), 267–283.

Dias, A. (2009). Technology enhanced learning and augmented reality: An application on multimedia interactive books. *International Business & Economics Review, 1*(1), 69–79.

Dietrich, N., Kentheswaran, K., Ahmadi, A., Teychené, J., Bessière, Y., Alfenore, S., Laborie, S., & Hébrard, G. (2020). Attempts, successes, and failures of distance learning in the time of COVID-19. *Journal of Chemical Education, 97*(9), 2448–2457.

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*, 287–312.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology, 18*(1), 7–22.

Ennis, R. H. (1991). Goals for a critical thinking curriculum. In A. Costa (Ed.), *Developing minds: A resource book for teaching thinking* (pp. 68–71). The Association of Supervision and Curriculum Development.

Ennis, R. H. (2011). *The nature of critical thinking: An outline of critical thinking dispositions and abilities.* Illinois College of Education. https://education.illinois.edu/docs/default-source/faculty-documents/robert-ennis/thenatureofcriticalthinking_51711_000.pdf

Erdogan, I., Ciftci, A., & Topcu, M. S. (2017). Examination of the questions used in science lessons and argumentation levels of students. *Journal of Baltic Science Education, 16*(6), 980–993.

Erduran, S. (2020). Science education in the era of a pandemic: How can history, philosophy and sociology of science contribute to education for understanding and solving the Covid-19 crisis? *Science & Education, 29*, 233–235.

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.

Erduran, S., Özdem, Y., & Park, J. Y. (2015). Research trends on argumentation in science education: A journal content analysis from 1998–2014. *International Journal of STEM Education, 2*(1), 1–12.
Erkens, G., & Janssen, J. (2006). Automatic coding of communication in collaboration protocols. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the 7th International Conference on Learning Sciences (ICLS)*, (pp. 1063–1064). International Society of the Learning Sciences. https://doi.org/10.5555/1150034
Faridi, H., Tuli, N., Mantri, A., Singh, G., & Gargrish, S. (2020). A framework utilizing augmented reality to improve critical thinking ability and learning gain of the students in Physics. *Computer Applications in Engineering Education, 29*, 258–273.
Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem based learning: The effects on learning achievement and attitude in physics education. *Computers & Education, 142*, 103635.
Fleck, S., & Simon, G. (2013, November). An augmented reality environment for astronomy learning in elementary grades: An exploratory study. In *Proceedings of the 25th Conference on l'Interaction Homme-Machine* (pp. 14–22). ACM.
Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review, 27*, 244–260.
Garzón, J., Pavón, J., & Baldiris, S. (2019). Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality*, 1–13
Garzón, J., Baldiris, S., Gutiérrez, J., & Pavón, J. (2020). How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. *Educational Research Review*, 100334.
Guba, E.G. & Lincoln, Y.S. (1989). *Fourth generation evaluation*. Sage Publications.
Hanid, M. F. A., Said, M. N. H. M., & Yahaya, N. (2020). Learning strategies using augmented reality technology in education: Meta-analysis. *Universal Journal of Educational Research, 8*(5A), 51–56.
Hsiao, H. S., Chang, C. S., Lin, C. Y., & Wang, Y. Z. (2016). Weather observers: A manipulative augmented reality system for weather simulations at home, in the classroom, and at a museum. *Interactive Learning Environments, 24*(1), 205–223.
Huang, K. T., Ball, C., Francis, J., Ratan, R., Boumis, J., & Fordham, J. (2019). Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications. *Cyberpsychology, Behavior, and Social Networking, 22*(2), 105–110.
Hwang, G. J., Wu, P. H., Chen, C. C., & Tu, N. T. (2016). Effects of an augmented reality-based educational game on students' learning achievements and attitudes in real-world observations. *Interactive Learning Environments, 24*(8), 1895–1906.
Ibáñez, M. B., Di Serio, Á., Villarán, D., & Kloos, C. D. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education, 71*, 1–13.
Jan, M. (2009). *Designing an augmented reality game-based curriculum for argumentation*. (Unpublished doctoral dissertation). University of Wisconsin-Madison.
Jermann, P., & Dillenbourg, P. (2003). Elaborating new arguments through a CSCL script. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 205–226). Springer.
Jiménez –Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 3–27). Springer.
Jiménez-Aleixandre, M. P., & Puig, B. (2012). Argumentation, evidence evaluation and critical thinking. In *Second international handbook of science education* (pp. 1001–1015). Springer, Dordrecht.
Ke, F., & Carafano, P. (2016). Collaborative science learning in an immersive flight simulation. *Computers & Education, 103*, 114–123.
Kennedy, G., Dalgarno, B., Bennett, S., Judd, T., Gray, K., & Chang, R. (2008). Immigrants and natives: Investigating differences between staff and students' use of technology. In *Hello! Where are you in the landscape of educational technology?. Proceedings Ascilite Melbourne, 2008*, 484–492.
Kerawalla, L., Luckin, R., Seljeflot, S., & Woolard, A. (2006). "Making it real": Exploring the potential of augmented reality for teaching primary school science. *Virtual Reality, 10*(3–4), 163–174.
Kirikkaya, E. B., & Başgül, M. Ş. (2019). The effect of the use of augmented reality applications on the academic success and motivation of 7th grade students. *Journal of Baltic Science Education, 18*(3), 362.
Kolsto, S. D. (2001). 'To trust or not to trust,…'-pupils' ways of judging information encountered in a socio-scientific issue. *International Journal of Science Education, 23*(9), 877–901.

Lin, S. S., & Mintzes, J. J. (2010). Learning argumentation skills through instruction in socioscientific issues: The effect of ability level. *International Journal of Science and Mathematics Education, 8*(6), 993–1017.

López-Faican, L., & Jaen, J. (2020). Emofindar: Evaluation of a mobile multiplayer augmented reality game for primary school children. *Computers & Education, 149*(2020), 103814.

Lu, S. J., Liu, Y. C., Chen, P. J., & Hsieh, M. R. (2020). Evaluation of AR embedded physical puzzle game on students' learning achievement and motivation on elementary natural science. *Interactive Learning Environments, 28*(4), 451–463.

Maykut, P., & Morehouse, R. (1994). *Beginning qualitative research: A philosophic and practical guide.* The Falmer Press.

Novak, A. M., & Treagust, D. F. (2017). Adjusting claims as new evidence emerges: Do students incorporate new evidence into their scientific explanations? *Journal of Research in Science Teaching, 55*(4), 526–549.

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology, 33*(3), 345–359.

OECD. (2003). *Literacy skills for the world of tomorrow: Further results from PISA 2003*. OECD Publishing.

Oestermeier, U., & Hesse, F. W. (2000). Verbal and visual causal arguments. *Cognition, 75*(1), 65–104.

Oleksiuk, V.P., Oleksiuk, O.R. (2020). Exploring the potential of augmented reality for teaching school computer science. In Burov, O.Yu., Kiv, A.E. (Eds.) *Proceedings of the 3rd International Workshop on Augmented Reality in Education (AREdu 2020)* (pp. 91–107).

Osborne, J., Erduran, S., & Simon, S. (2004). Enhancing the quality of argument in school science. *Journal of Research in Science Teaching, 41*(10), 994–1020.

Parker, J., & Heywood, D. (1998). The earth and beyond: Developing primary teachers' understanding of basic astronomical events. *International Journal of Science Education, 20*, 503–520.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage.

Pellas, N., Fotaris, P., Kazanidis, I., & Wells, D. (2019). Augmenting the learning experience in primary and secondary school education: A systematic review of recent trends in augmented reality game-based learning. *Virtual Reality, 23*(4), 329–346.

Ploman, E. W. (1987), Global learning: A challenge. In C. A., Taylor (Ed.) *Science education and information transfer* (pp. 75–80). Oxford: Pergamon (for ICSU Press).

Radu, I., & Schneider, B. (2019). What can we learn from augmented reality (AR)? Benefits and drawbacks of AR for inquiry-based learning of physics. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12).

Rideout, V. J., Foehr, U. G., & Roberts, D. F. (2010). *Generation M [superscript 2]: Media in the lives of 8-to 18-year-olds*. Kaiser Family Foundation.

Rieber, L. P., & Kini, A. S. (1991). Theoretical foundations of instructional applications of computer-generated animated visuals. *Journal of Computer-Based Instruction, 18*, 83e88.

Romano, M., Díaz, P., & Aedo, I. (2020). Empowering teachers to create augmented reality experiences: The effects on the educational experience. *Interactive Learning Environments*, 1–18.

Sadler, T. D., & Donnelly, L. A. (2006). Socioscientific argumentation: The effects of content knowledge and morality. *International Journal of Science Education, 28*(12), 1463–1488.

Sahin, D., & Yilmaz, R. M. (2020). The effect of augmented reality technology on middle school students' achievements and attitudes towards science education. *Computers & Education, 144*(2020), 103710.

Sampson, V., & Clark, D. B. (2011). A comparison of the collaborative scientific argumentation practices of two high and two low performing groups. *Research in Science Education, 41*(1), 63–97.

Sampson, V., & Gleim, L. (2009). Argument-driven ınquiry to promote the understanding of important concepts & practices in biology. *The American Biology Teacher, 71*(8), 465–472.

Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-driven ınquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education, 95*(2), 217–257.

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.

Saribas, D., & Çetinkaya, E. (2021). Pre-service teachers' analysis of claims about COVID-19 in an online course. *Science & Education, 30*(2), 235–266.

Schnotz, W., & Lowe, R. K. (2008). A unified view of learning from animated and static graphics. In R. K. Lowe & W. Schnotz (Eds.), *Learning with animation: Research implications for design* (pp. 304–357). Cambridge University Press.

Setozaki, N., Suzuki, K., Iwasaki, T., & Morita, Y. (2017). Development and evaluation of the usefulness of collaborative learning on the tangible AR learning equipment for astronomy education. *Educational Technology Research, 40*(1), 71–83.

Shelton, B. E. & Hedley, N. R. (2002). Using augmented reality for teaching earth-sun relationships to undergraduate geography students. *In Augmented Reality Toolkit, The First IEEE International Workshop* (8).

Sin, A. K., & Zaman, H. B. (2010). Live solar system (LSS): Evaluation of an augmented reality book-based educational tool. *In Proceedings of 2010 International Symposium on Information Technology, 1*, 1–6.

Squire, K. (2006). From content to context: Videogames as designed experience. *Educational Researcher, 35*(8), 19–29.

Squire, K. D., & Jan, M. (2007). Mad City Mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of Science Education and Technology, 16*(1), 5–29.

Stark, R., Puhl, T., & Krause, U. M. (2009). Improving scientific argumentation skills by a problem-based learning environment: Effects of an elaboration tool and relevance of student characteristics. *Evaluation & Research in Education, 22*(1), 51–68.

Syawaludin, A., Gunarhadi, R., & P. (2019). Development of augmented reality-based interactive multimedia to improve critical thinking skills in science learning. *International Journal of Instruction, 12*(4), 331–344.

Taylor, C. A. (1987). *Science education and information transfer* (pp.1–15). Oxford: Pergamon (for ICSU Press)

Toulmin, S. E. (1990). *The uses of argument* (10th ed.). Cambridge University Press.

Venville, G. J., & Dawson, V. M. (2010). The impact of a classroom intervention on grade 10 students' argumentation skills, informal reasoning, and conceptual understanding of science. *Journal of Research in Science Teaching, 47*(8), 952–977.

Virata, R. O., & Castro, J. D. L. (2019). Augmented reality in science classroom: Perceived effects in education, visualization and information processing. In *Proceedings of the 10th International Conference on E-Education, E-Business, E-Management and E-Learning* (pp. 85–92).

Von Aufschnaiter, C., Erduran, S., Osborne, J., & Simon, S. (2008). Arguing to learn and learning to argue: Case studies of how students' argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching, 45*(1), 101–131.

Wu, Y. T., & Tsai, C. C. (2007). High school students' informal reasoning on a socio-scientific issue: Qualitative and quantitative analyses. *International Journal of Science Education, 29*(9), 1163–1187.

Yen, J. C., Tsai, C. H., & Wu, M. (2013). Augmented reality in the higher education: Students' science concept learning and academic achievement in astronomy. *Procedia-Social and Behavioral Sciences, 103*, 165–173.

Yildirim, I., & Seckin-Kapucu, M. (2020). The effect of augmented reality applications in science education on academic achievement and retention of 6th grade students. *Journal of Education in Science Environment and Health, 7*(1), 56–71.

Yu, K. C., & Sahami, K. (2007). Visuospatial astronomy education in immersive digital planetariums. *Communicating Astronomy with the Public*, 242–245.

Yuen, S., Yaoyuneyong, G., & Johnson, E. (2011). Augmented reality: An overview and five directions for AR in education. *Journal of Educational Technology Development and Exchange, 4*(1), 119–140.

Zhang, J., Sung, Y.-T., Hou, H.-T., & Chang, K.-E. (2014). The development and evaluation of an augmented reality-based armillary sphere for astronomical observation instruction. *Computers & Education, 73*, 178–188.

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.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.