ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 25, NO. 2, 68-93

ICRSME
International Consortium for Research in
Science & Mathematics Education

The Structural and Contextual Quality of Preservice Elementary Teachers' Argumentative Discussions

Evren Cappellaro
Akdeniz University

Demet Seban
Alanya Alaaddin Keykubat University

ABSTRACT

This study investigates the preservice elementary teachers' ability to use the components of an argument structure on a given science topic as well as the quality of the developed argument during their sophomore and senior year. In doing so, the study also aimed to discuss the effect of teacher education programs on the development of argumentation skills of students over two years. A qualitative research approach was applied. The data were collected from the same participants at the end of their second and fourth years through having them discuss different generic frameworks. The argumentative discourses were analyzed in four stages. The scientific argumentations' overall quality was at low-level. The participants' claims were generally established without using data and warrants. They failed to transfer their scientific knowledge into discussion using qualifiers and rebuttals. The quality of the data, claim, and warrants were insufficient and included misconceptions. The teacher education program increased the students' scientific knowledge to some extent. Courses with more argumentation-based discussions could help students to increase the quality of their argument and their capabilities to use components of an argument structure.

Keywords: preservice elementary teachers, argumentation, science education

Introduction

Today, the obtaining of first-hand information has become significantly easier. This development has placed emphasis on acquisition of scientific thinking skills rather than teaching of scientific knowledge and concepts in science education. Now, students are expected to use scientific knowledge, to understand its impact on their lives, to achieve conceptual learning through conducting studies and experiments, and subsequently to discuss and share what they have learned (Osborne et al., 2001). Increasing emphasis is placed on the teaching of argumentation-based science among the many other theories and approaches that have been developed in order to gain these skills. In recent years, the attention paid to the research in this field has been growing continuously (Erduran et al., 2015; Lee et al., 2009; Lin et al., 2014).

All forms of discussion enable students to acquire skills for information collection, selection, and inquiry and ensure their engagement in the social structuring of knowledge (Kuhn, 2005). However, argumentation differs in the process of transposition and hypothesis on learning. The importance of this approach comes from its practical aspects reflecting the rational thinking processes of scientists. That is why science learning requires argumentation skills as well as scientific knowledge. Argumentation-based scientific education facilitates meaningful learning for students. It helps students use cognitive/metacognitive strategies and processes, develops their communication skills, supports their critical thinking skills, promotes scientific literacy, and makes it easier for them to understand scientific culture and practice (Jimenez-Aleixandre & Erduran, 2008).

Teaching argumentation in science education also conveys an adequate image of science through constructing and analyzing arguments related to social applications and implications of science (Driver et al., 2000). This way, students can also realize that science and scientific knowledge is questioned, discussed, and subject to change. Argumentation does not only improve scientific thinking and inquiry skills but also the way students approach to the nature of science (Duschl & Osborne, 2002; Tumay & Koseoglu, 2011). Introducing students to this approach at an early age also contributes to their understanding of the nature of science and creates a positive impact on their approach to scientific concepts and sciences in future study (Ozdem et al., 2013).

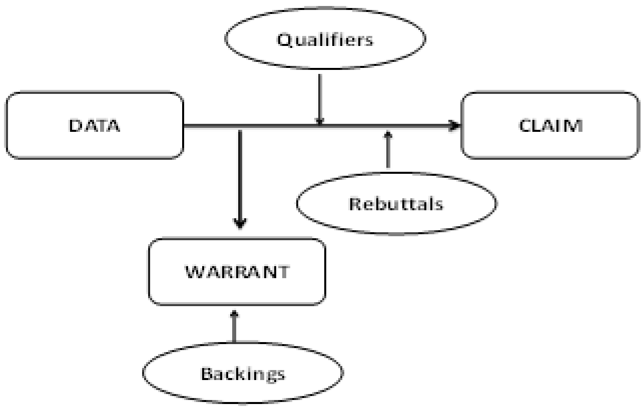
The legitimate answer to assess quality of arguments relies on the definition of argument and argumentation theories. Argumentation involves diverse meanings and many scientists have contributed to defining the concept of argument in the related literature in various ways (Jimenez-Aleixandre & Erduran, 2008; Kuhn, 1992; Plantin, 2005). Jimenez-Aleixandre and Erduran (2008) emphasized that the structuring of knowledge in the sciences is linked to evidence-based justification and that the claims should be presented within a logical context through its correlation to the data and evidence obtained from different sources. Therefore, they defined argument as constructing a connection between data and claims by means of rebuttal or evaluation in light of theoretical or empirical data in regard to a scientific topic/knowledge.

Toulmin made significant contributions to the development of the definition of scientific argumentation as a concept. He explains the structure of argument in a model consisting of four main elements (claim, data, warrant, backing) and two auxiliary elements (qualifiers, rebuttal) (Toulmin, 1958). According to this model, the knowledge and facts to support the claim is defined as *data*; the conclusion that is drawn is defined as *claim*; causes and principles to validate the relation between the data and claim is defined as *warrant*; and assumptions agreed to validate certain justifications is defined as *backing*. For more complex arguments, Toulmin then defined two more elements, namely qualifiers and rebuttals. *Qualifiers* are statements that limit the conditions under which the claim is true. *Rebuttals* are counterarguments or statements indicating circumstances when the claim is not true.

In examining the forms of argument, Walton (1996) defines argumentation schemes as abstract structures and frames them as dialogical. Many different argumentation schemes were distinguished based on their forms and content. Every argumentative scheme provides a set of questions for different types of reasoning such as inductive, deductive, defaultive, or abductive. They represent the structures of common types of arguments used in different contexts like everyday, legal, and scientific discourse. To analyze an argument, the argumentation scheme needs to be identified first. If the scheme is identified, the explicit and implicit features of the premises or the relation between premises and conclusion can be identified (Walton et al., 2008). Anderson et al. (2001) also point to the dialogical nature of the process of argumentation and emphasize the assumption of reasoning and representing contrasting perspectives. They argue that extended arguments can be broken down into recurrent patterns which they call "argument stratagems" (p. 2). A complete argument stratagem should include five different categories of information: the purpose, the conditions, the forms, the consequences, and the objections. Each of these stratagems depends on an argument schema that helps to organize information, retrieve relevant information from memory, facilitate argument invention, provide the basis for anticipating objections, and find flaws in the arguments.

According to Sampson and Clark (2008), a great number of diverse analytical and conceptual frameworks are used to analyze argumentation processes carried out in research done in relation to argumentation-based education in science education. According to the researchers, the models developed by Toulmin (1958) and Schwarz et al. (2003) were two different approaches employed in research investigating the arguments constructed by students. The common aspect of these two approaches is that they can easily be generalized and adapted to diverse research fields. Toulmin's (1958) argument pattern is one of the most common argument models employed (Figure 1). In this model, what is investigated is usually the argument elements and not the content of an argument.

Figure 1
Toulmin Model of Argument



According to Toulmin's model, arguments that are deemed as strong can sometimes fall short of scientific content. As the presence of argument components cannot be an indicator of a qualified and well-established argument, Toulmin's model is used together with other supporting approaches when conducting research (Sampson & Clark, 2008). The framework of Schwarz et al. (2003), on the other hand, is for examination of arguments that have been developed in written discourse. This framework is used mainly for analyzing the argument in terms of content.

The discussions on how argumentation is situated in science and how it intervenes in science education legitimized argumentation-based science education. Different integration of argumentation as a part of curriculum and teacher training for the implementation of argumentation in science education became two aspects to consider for researchers who work in this area. Most national curricula expect students to acquire general argumentative skills, and science educational policies across the world highlight argumentation as a process of scientific knowledge construction (Erduran & Jimenex-Aleixandre, 2008). In Turkey, there were no standards directly related to argumentation in elementary school programs until 2015. "An interdisciplinary perspective based on an inquiry-based learning approach" (MEB, 2015) was adopted and argumentation-based science teaching was explicitly highlighted in the science curriculum beginning in elementary grades. Adopting a new revised curriculum requires time, and the success of implementation depends on the teachers' ability to implement argumentation practices. Besides providing in-service teachers with the necessary training, teacher education programs should consider new developments in the curriculum and revise their own curricular activities according to these changes. In addition to subject matter knowledge, preservice teachers are also educated on different instructional approaches and teaching techniques. The instruction that involves higher order thinking should be specifically designed to actively engage students to learn the process of thinking as scientific argumentation skills do not naturally develop and argumentation is not a process that spontaneously transpires in the class (McNeill & Pimentel, 2010).

The majority of studies carried out in relation to the use of argumentation in science education were done on preservice science teachers, inservice teachers at the level of middle school and high school, or students studying at this level of education (Cavagnetto, 2010). The number of studies on primary school students and teachers, on the other hand, is scarce in comparison to these studies (Cavagnetto et al., 2010; Kim & Hand, 2015). Therefore, preservice teachers who teach science should be equipped in terms of argumentation-based teaching. The effective use of this approach by the elementary level preservice teachers who are to guide students in argumentation-based science teaching and encourage meaningful learning depends on the fact that they understand the nature of scientific knowledge and that they have the necessary knowledge and skills for implementing this approach. This is also a necessity because success in middle school depends on how well students were introduced to argumentation-based reasoning and discussion.

Given this theoretical and empirical background to the problem, this study concerns preservice elementary teachers' scientific argumentation development from the second to fourth year of their program. In doing so, the study also aimed to discuss the effect of teacher education programs on the development of argumentation skills. Therefore, the research questions are:

- What are the components of argument structures of scientific discussions developed by sophomore and senior preservice elementary teachers?

- What is the quality of arguments developed by sophomore and senior preservice elementary teachers?

- To what extent does the structural and contextual quality of preservice elementary teachers' scientific argument develop from sophomore to senior year?

Method

This study aims to examine preservice elementary teachers' ability to use argumentation structure and the quality of the developed argument content. It also has the objective to discuss the contribution made by the education received over two years to the structural and contextual quality of preservice elementary teachers' scientific argument. The study employs a qualitative research method as it provides a holistic picture of the phenomena to understand the research problem (Cresswell, 2007). Qualitative research aims to draw a comprehensive picture and to interpret the meanings inferred from the phenomenon (Denzin & Lincoln, 1998).

Participants

The study was administered to 33 preservice elementary teachers. The participants are the same students who studied during the academic year of 2013-2014 in their second year and the academic year of 2015-2016 in their fourth. Seventeen female and 16 male students were voluntarily involved in the study. Purposeful sampling techniques were used to select participants in which the researcher selected participants who were associated with the research problem being studied (Creswell, 2007). The inclusion criteria were their success in science lessons in the first and second years of the education program.

Data Collection

With a view to examine the preservice elementary teachers' argumentation-based scientific discussions and the change that occurred in the structure and the content of these arguments, they were given a separate appointment at the end of the final exams so as to collect the data on two different occasions. The first data collection was carried out after the completion of year two by the students. Year two was chosen because the preservice teachers completed General Biology, General Physics, Chemistry, Instructional Principles and Methods, Science and Technology Laboratory Practices I and II at the end of that academic year. The students are expected to be knowledgeable about the basic concepts of science at the end of this year. Two years later, the second data collection was conducted right before the participants graduated. The students completed Science and Technology Education I and II, Life and Social Studies Education, and School Experience and Teaching Practices I and II throughout the last two years of the program. The preservice teachers gained experience as to how to convey knowledge and skills with regard to the science content they had learned. The students did not get any specific intervention program, nor were they trained specifically about argumentation using Toulmin's argumentation model. This was not preferred because teacher education programs are expected to provide students with necessary knowledge that they will need to make arguments.

Five groups of preservice teachers were formed, and each group received one generic question. When the groups were established, partner selection choices of the participants were taken into account and a group consisted of at least six students.

Data Collection Tool

The questions distributed to the groups were selected from among the questions developed by Osborne et al. (2001) with the contributions of a group of teachers. The researchers designated various argument types and gave exemplary discussion questions for each type of argument to improve the scientific argumentation skills of children. The questions were translated from English to Turkish and used after being revised in line with the opinions of two expert science teachers (Table 1).

Data Analysis Method

The group discussions were audio-recorded and lasted about 20 minutes. The discussion transcripts were analyzed in four stages. During the analysis process, categories that were pre-determined were used for each stage.

In the first stage, all data were analyzed by using Toulmin's (1958) argument pattern. The dimensions constituting this model are claim, data supporting the claim, warrants indicating the relation between the data and the claim, backing strengthening the warrants, qualifiers, and finally the rebuttals showing the conditions or events under which the claim is invalid. This stage of the study was conducted to determine which dimensions of the argument were used by the preservice teachers. The analysis of the data was undertaken by two researchers. One of the researchers has Ph.D. in science education and the other in Curriculum and Instruction in elementary education. One researcher is an expert in teaching elementary and middle school science as the other works within literacy and is very familiar with teaching how to write argumentative essays. General evaluations were noted down after having listened to the interview records. Subsequently, the researchers identified the claims, warrants, backings, qualifiers, and finally assumptions used through scrutinizing the written texts and then checked whether the discourses were coded into similar categories. The discourses that were not coded similarly were discussed and evaluated under the related category upon agreement.

In the second stage, the content of the argument was evaluated in terms of strength and weakness. In this evaluation, the framework developed by Okada and Buckingham Shum (2008) was used (Table 2). In this stage, the data analysis was conducted without considering the qualities of argument elements, only taking into account their presence and absence. In the tables, "2Y" represents second year students, "4Y" represents fourth year students. The subsequent number represents the group and the letter represents the name assigned to the preservice teacher in that group.

Table 1
Materials for Argumentation

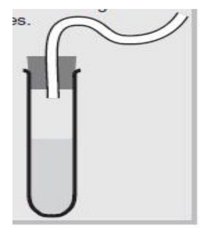
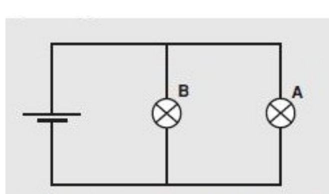
The Group	Generic Frameworks
1.	**Competing theories** **Theory 1**: Microbes are made from rotting material. This is helped by flashes of lightning, bright light and warmth. **Theory 2**: Microbes are carried in the air, probably on dust, and cannot be made out of dead matter. *Discuss whether the following pieces of evidence support Theory 1, Theory 2, both or neither.* **a.** Boiled soup in a sealed glass flask will keep forever. **b.** If cartons of milk are opened, they will not stay fresh. **c.** If the air was full of microbes, you wouldn't be able to see at all. **d.** If milk is heated and sealed, it will keep for several days. **e.** Boiling any material kills the vital ingredients needed to make microbes. **f.** Boiled soup, exposed to the air with a special S bend tube in it like the one shown here, does not go off. **g.** Food goes off more quickly in the summer when it is warm and humid. 
2.	**Constructing an argument** Heavier things do not always fall faster. *Look at the following statements of evidence. Discuss them with the others in your group and put them in a logical order to justify the statement above.* **a.** A penny and a brick reach the ground at the same time when dropped from the same height. **b.** Air resistance is a force which opposes motion. **c.** All things fall at the same rate if you ignore air resistance. **d.** A piece of paper falls much more slowly than a brick.
3.	**Understanding an argument** *Which of the following arguments provide good evidence that matter is made up of particles, and why?* **a.** Air in a syringe can be squeezed. **b.** All the crystals of any pure substance have the same shape. **c.** Water in a puddle disappears. **d.** Paper can be torn into very small pieces.
4.	**Experimental data** Everybody in the class measured the boiling point of water. They obtained the following results. 96 ˚C, 94 ˚C, 102 ˚C, 106 ˚C, 108 ˚C, 92 ˚C, 101 ˚C, 86 ˚C, 97 ˚C, 103 ˚C *In your group discuss:* **a.** Why might they disagree? **b.** How might they agree on a value?
5.	**Predicting, observing and explaining** Bulb A and Bulb B are two identical bulbs. *What will happen to the brightness of lamp B when lamp A is unscrewed?* *Discuss in your group and give reasons for what you think will happen.* 

Table 1
Categories of Argument Content

Argument Level	Statement	Example
Very weak	Only claim	*I think the pencil would fall faster.* (2Y-2C)
Weak	Claim and warrants (usually based on beliefs)	*As the density of air is close to that of the paper, the fall would slow down at this point.* (2Y-2D)
Moderate	Claim, (weak) warrants, data, or rebuttals	*Volume has effect. If the volume is bigger, then the frictional force applies more power.* (2Y-2A)
Strong	Claim, warrants, rebuttals, and/or data	*The volume of brick is bigger. Of course, if it is a full brick. As the volume of the brick is bigger, more frictional force applies.* (2Y-2C)

In the third stage, the claims of the arguments were evaluated as being right, wrong, or having misconceptions, whereas the warrants and data were evaluated as being sufficient, insufficient, and having misconceptions (Table 3).

Table 2
Categories of Argument Quality

Categories	Aspect	Example
Right claim	Scientifically right	*I think this can stay as it is forever because it has no contact with air. (4Y-1D)*
Wrong claim	Scientifically wrong	*I think the pencil would fall faster.* (4Y-2C)
Claim with misconception	Partially scientifically right, but it has misconceptions	*It is not in liquid phase at 103 degrees centigrade.* (2Y-4B)
Sufficient warrants	It supports the claim in every aspect.	*I begin first. Water, pure water boils at 100 degrees centigrade under atmospheric pressure. The characteristics of water, its purity, and its location and therefore its atmospheric pressure, they all change the boiling point. So, the differences in measurements may be due to these. The purity of water may be disrupted, changed or the atmospheric pressure where the boiling point is being measured may be increased or decreased. Thus, for the option A, I believe the different measurements may be the consequences of these changes.* (4Y-4E)
Insufficient warrants	It supports the claim in every aspect but is scientifically weak.	*As you know, the air in it can be compressed and it has particles inside, so that is why we can compress it.* (2Y-4B)
Warrants with misconception	It supports the claim partially but has misconceptions.	*The speed of putrefaction decreases related to working of enzymes.* (2Y-1A)
Insufficient data	Data supporting the claim only from one aspect.	*Ultimately, the matter itself is made of atoms, isn't it?...* (2Y-3B)
Data with misconception	Data having misconceptions.	… *These particles should be flexible or should shrink.* …(4Y-3F)

In the fourth stage, the data collected at the end of the second and fourth year were compared considering the aforementioned dimensions.

Results

In this phase, the arguments constructed by all the preservice teachers who participated in the research were first evaluated in terms of argument components and then their qualities. The groups were compared overall under these two elements.

Basic Components of Constructed Scientific Arguments

The arguments structured by the preservice teachers engaged in the research during the discussions were analyzed according to the aforementioned evaluation criteria. When the arguments are compared as frequency values, 127 arguments were developed in total at the end of year two, whereas this number increased to 139 at the end of year four. The highest number of arguments developed at the end of year two belonged to the second group discussing free fall. Meanwhile, the groups in year four developed almost equal numbers of arguments in general. The fifth group developed the lowest number of arguments when asked to construct arguments in relation to electric circuits at the end of years two and four (Table 4).

Table 4

Percentage Distribution of Argument Components in Group Discussions

Categories	Groups											
	1		2		3		4		5		Total	
	2Y	4Y	2Y	4Y	2Y	4Y	2Y	4Y	2Y	4Y	2Y	4Y
	%	%	%	%	%	%	%	%	%	%	%	%
Very weak	5.5	7.1	14.9	5.7	5.5	10	13.3	5	0.7	0	40	28
Weak	7.8	12.2	11	6.4	7.8	5	3.9	10	2.3	4.3	33	38
Moderate	3.9	1.4	7.8	10.7	6.2	8.6	1.5	4.3	3.9	4.3	24	30
Strong	0	0	2.3	0	0.7	0.7	0	2.1	0.7	0.7	3	4
Total	17.3	20,7	36	22.8	20.2	24.3	18.7	21.4	7.6	9.3	100	100

When Table 4 is examined, according to Toulmin's argumentation model, it was observed that the majority of the arguments developed at the end of year two involved arguments that were very weak and weak, and at the end of year four, weak and moderate. While very weak arguments were as high as 40% at the end of year two, this number decreased to the level of 28% at the end of year four. Moderate level arguments were equal to 25% of all the arguments developed at the end of year two, while this increased to 30% at the end of year four. The number of strong level arguments were similar at the end of years two and four.

When the groups were compared according to the argument categories, it was seen that the highest number of very weak arguments belonged to the second and fourth groups at the end of year two. While the rate of arguments of these groups under the very weak category decreased at the end of year two, this rate increased for the first and third groups. When, however, the weak arguments composed of claims and warrants developed in all the groups were compared, it was concluded that at the end of year two, the second group created more weak arguments in comparison with other groups, and at the end of year four, the first group created more weak arguments in comparison with other groups. As for the fifth group, they produced a smaller number of arguments in this category as compared to other groups both at the end of year two and year four.

The highest number of moderate level arguments was produced by the second group. When all the arguments at the moderate level were compared, the rate of these arguments produced by the first group decreased at the end of year four, whereas it increased in other groups. It was seen that the first and fourth groups did not form any strong arguments at the end of year two when the strong arguments developed by all the groups were examined, whereas the second group created a higher number of arguments in this category. At the end of year four, no arguments were detected in this category, either in the first or second group. While there were arguments present in this category in the discussion carried out by the second group at the end of year two, they were not able to produce any strong arguments at the end of year four. Although the fourth group did not make any strong arguments at the end of year two, they did produce strong arguments at the end of year four. As for the other groups, the rate of arguments in this category did not change.

Understanding the Nature of Scientific Argument: Competing Theories

In this category, it was observed that most of the participants discussing two conflicting theories in relation to microbes developed very weak and weak arguments at the end of years two and four (Table 5). The rate of weak arguments increased at the end of year four. The arguments with claims and rebuttals were present at the end of year two, while there were none at the end of year four. In comparison to year two, year four arguments comprised all of the following: claim, warrant and rebuttal. As for the moderate level arguments in this category, while the arguments in which claim and rebuttals stand out at the end of year two, no arguments in this category were observed to have been formed at the end of year four.

Table 5
Percentage and Frequency Distribution of Understanding the Nature of Argument Components

		2Y		4Y	
	Categories	f	%	f	%
Very weak	Claim	7	32	10	33
Weak	Claim+ warrant	10	46	17	57
Moderate	Claim+ data	1	5		
	Claim+ rebuttals	3	14		
	Claim+ data+ warrant	1	5	2	7
	Rebuttal				
Strong	Claim+ warrant+ rebuttal			1	3
	Claim+ data+ warrant+ rebuttal				
Total		22	100	30	100

The data used in the discussion is the data that were created based on the information which was embedded in the question itself. While only one participant used data in the arguments at the end of year two, two different participants made use of warrants and claims in addition to data at the end of year four. Furthermore, another participant developed an argument that contained claim, warrant, and rebuttal:

> *I think this can stay as it is forever as it has no contact with the air because microbes are transmitted through air. However, if air is present or leaks in the ambient, then it may not endure.* (4Y-1D)

Constructing an Argument: Free Fall

When participants constructed arguments on the claim that heavy objects do not always fall fast were analyzed, it was observed that while very weak arguments were formed at the end of year two, it was the moderate-level arguments with the highest number obtained at the end of year four (Table 6). The rate of very weak arguments decreased at the end of year four. When the category with moderate arguments is examined, what is striking is that the participants in this group used rebuttals or data together with the claim. While no arguments were detected having claim, data, and warrant at the end of year two, such arguments in this category had the highest number at the end of year four. Strong arguments were identified at the end of year two; however, the participants did not construct any strong arguments at the end of year four. There were participants in the group who could not form any arguments at all.

Table 6

Percentage and Frequency Distribution of Argument Components for Constructing an Argument

		2Y		4Y	
	Categories	f	%	f	%
Very weak	Claim	19	42	8	25
Weak	Claim+ warrant	14	31	9	28
Moderate	Claim+ data	2	4	2	6
	Claim+ rebuttals	3	7	4	13
	Claim+ data+ warrant			6	19
	Rebuttal	5	11	3	9
Strong	Claim+ warrant+ rebuttal	1	2		
	Claim+ data+ warrant+ rebuttal	1	2		
Total		45	100	32	100

Although during the discussion the participants mentioned in their arguments the concept of mass in relation to free fall in general, they included too many unrelated concepts such as density, volume, and weight. Moreover, not one among the participants used the most important concepts of free fall; that is, gravitational acceleration, cross-section, air resistance, or friction:

The one with more mass would fall faster. (2Y-1D)

Let's not forget that its density is different. (4Y-1A)

Understanding an Argument: Structure of Particulate Matter

When Table 7 is examined, it can be seen that the highest numbers of argument categories for the four claims about the particulate matter's structure were weak and moderate arguments at the end of year two and very weak and moderate arguments at the end of year four. While weak-level arguments were lesser in the group discussion at the end of year four, very weak arguments increased in number. Meanwhile the rate obtained for strong arguments showed no change. The claims were mostly presented with data in the moderate arguments at the end of year two. At the end of year four, in addition to arguments in which claim and data were used together, arguments with claim, data, and warrants were detected:

We know from the empirical evidence that syringe can be pressurized, in other words, as the distance of particles in gas phase is larger, they can be compressed. (2Y-3B)

We are saying that this would apply for every matter, right? But, let's consider pebbles instead of sand. Pebbles can, too, accumulate and when it does and when we pour water in it, the water would flow through the pebbles. They would not hold anything. (4Y-3E)

Table 7

Percentage and Frequency Distribution of Argument Components for Understanding an Argument

		2Y		4Y	
	Categories	f	%	f	%
Very weak	Claim	7	27	14	41
Weak	Claim+ warrant	10	39	7	20
Moderate	Claim+ data	4	15	5	15
	Claim+ rebuttals	2	8	3	9
	Claim+ data+ warrant	2	8	4	12
	Rebuttal				
Strong	Claim+ warrant+ rebuttal	1	4	1	3
	Claim+ data+ warrant+ rebuttal				
Total		26	100	34	100

Albeit similar to the arguments in the discussion, the only strong argument identified at the end of years two and four involved claim, warrant, and rebuttal. When examined scientifically, the warrant supported the claim from only one aspect. For instance, the most striking characteristics of the strong argument presented at the end of year two was that no scientific terminology was used except for the one provided in the question:

If it was not a particulate matter, if water was in bulk, it could not get through the soil or evaporate into the air as a whole; thus, this shows that it has a particulate structure. (2Y-3E)

The group members failed to present counter-claims or did not address the validity of the claims as they did in other questions. The claims lacked scientific concepts and usually focused on options A and C. Although one participant was curious about whether the particles would differ for option B, s/he failed to turn it into a claim. And the question asked by this participant was not answered by other participants. This was also the case when other options were addressed during the discussion process. Instead of providing an answer to the claims presented, the participants sometimes changed the subject at hand, which created an environment in which the group conducted a disorganized conversation-like session. In this group's discussion, the scientific concepts were mentioned much less than the previous group.

Interpreting Experimental Data: Boiling Point of Water

The fourth group was asked to elaborate on the reasons for different results obtained from an experiment in which water's boiling point was measured. When this discussion was examined, it was seen that very weak and weak arguments were produced at the end of year two, and weak and moderate arguments were formed at the end of year four. While, the number of very weak arguments lessened at the end of year four, the number of moderate arguments increased. No strong arguments were detected at the end of year two, whereas at the end of year four, strong arguments were present (Table 8).

Table 8

Percentage and Frequency Distribution of Argument Components for Interpreting Experimental Data

	Categories	2Y		4Y	
		f	%	f	%
Very weak	Claim	17	71	7	23
Weak	Claim+ warrant	5	21	14	47
Moderate	Claim+ data			3	10
	Claim+ rebuttals	2	8	1	3
	Claim+ data+ warrant			2	7
	Rebuttal				
Strong	Claim+ warrant+ rebuttal			2	7
	Claim+ data+ warrant+ rebuttal			1	3
Total		24	100	30	100

The highest number of very weak arguments produced among groups was for this question; that is, discussion where empirical data were to be interpreted in relation to the boiling point of water. The participants made consecutive claims during the discussion:

> *The water could be calcareous or salty … could be a student error.* (2Y-4A)

When the group discussion was analyzed, it was observed that the participants did not hold a discussion on the impact of the claims presented on the data in the question. When, however, the claims were examined, as opposed to the structures of the claims presented by other groups, it was seen that the participants in this group did not use simple present tense, rather they used modal verbs of probability:

> *It could be fresh water.* (4Y-4B)

The group focused on a mixture of student error or measurement error rather than the concepts of height and pressure. As for option B in the question, at the end of year two, they recommended calculating the arithmetic average after summing up all related values. Only one participant (4D) objected to this and recommended repeating the experiment under a supervisor but the calculation of arithmetic average was approved by the other participants. However, when this claim was presented again at the end of year four, it was rejected by other participants:

> *When we discussed this the last time, I remember very well that we had said that we could calculate the arithmetic average of all the values.* (4Y-4C)

Predicting and Explaining: Parallel Electric Circuit

The lowest number of arguments constructed among the groups was the discussion where the group addressed the results of loosening the connection of two parallel bulbs attached to an electric circuit (Table 9). Although very few arguments were formed, they were superior in strength when compared to those created by other groups. While the moderate arguments were almost equal to half of the total number of arguments, they usually composed claim, data, and warrant. The strong arguments constructed at the end of year two included elements of data, warrant, and rebuttal:

> *I believe no changes will be observed because they are both identical and the same current flows through both. If they were switched on, they would glow in the exact way. They would not be brighter than one another. Loosening A would not mean to increase the brightness of B.* (2Y-5D)

Table 9
Percentage and Frequency Distribution of Argument Components for Predicting and Explaining

		2Y		4Y	
		f	%	f	%
Very weak	Claim	1	10		
Weak	Claim+ warrant	3	30	6	46
Moderate	Claim+ data	2	20	2	15
	Claim+ rebuttals				
	Claim+ data+ warrant	3	30	4	31
	Rebuttal				
Strong	Claim+ warrant+ rebuttal			1	8
	Claim+ data+ warrant+ rebuttal	1	10		
Total		10	100	13	100

At the end of year four, data were included in the strong arguments:

> *I think it would not change because the current flowing through both parallel identical bulbs is the same. For instance, the bulbs at home are connected in parallel and when one is switched on, the other's brightness does not dim, right? But, if they were connected in series, then its brightness would dim.* (4Y-5C)

Among the group discussions considered to be the shortest was the one conducted for this question. The participants in this group carried out a discussion where no one undertook the effort to refute or support each other's ideas, although many ideas were brought forward, which was the case for the third group as well. Each member of the group participated in the discussion equally. Concepts of voltage and impedance used for the explanation of electric circuits were not mentioned in the discussion.

Basic Qualities of Constructed Scientific Arguments

Another objective of the research was to examine the qualities of argument elements developed by the participants at the end of years two and four (Table 10). Numbers of claim, warrants, and data were examined at the end of years two and four, respectively: 122 and 136 claims; 50 and 76 warrants; and 13 and 30 data. The highest number of claims was obtained in the second group's discussion on free fall with 40 and 36 claims at the end of years two and four, respectively. While the second group included the highest number of warrants (16 warrants) in their discussion at the end of year two, it was the first group that reached the highest number of warrants with 19 warrants at the end of year four. The arguments with the highest number of data were made by the participants in the third group. The first and fifth groups used the least of data by using data in only one argument.

When Table 10 is considered overall, it can be deducted that the preservice teachers at the end of year four made claims that were scientifically more accurate and their misconceptions lessened; that a sufficient number of warrants, although not many, was present and the misconceptions in warrants lessened; and that adequate level of data were included.

When their choice of claim was examined, it can be stated that while some of the claims were right, some of them were wrong and some of them had misconceptions. The second group who produced the highest number of claims presented more scientifically wrong claims with more misconceptions as opposed to other groups. However, overall, they produced correct claims at the end of years two and four. Only the fifth group at the end of year two presented claims of which 80 percent was wrong scientifically:

> *I believe that since the bulbs are connected in parallel, loosening one's connection to the electric circuit would increase the brightness of the other bulb.* (2Y-5A)

Table 10
The Qualities of the Argument Components in the Group Discussions

		Groups											
		1		2		3		4		5		Total	
		2Y	4Y	2Y	4Y	2Y	4Y	2Y	4Y	2Y	4Y	2Y	4Y
Categories		%	%	%	%	%	%	%	%	%	%	%	%
Claim	Right	72	93	47	48	78	84	83	93	20	85	64	82
	Wrong	22	7	2	38	5	0	11	0	80	5	22	11
	Misconception	4	0	27	14	17	6	6	7	0	0	14	7
Warrant	Sufficient	0	0	0	0	0	0	0	11	0	0	0	3
	Insufficient	90	100	31	40	56	83	100	74	71	82	64	75
	Misconception	10	0	68	60	44	17	0	15	28	8	36	22
Data	Sufficient	0	0	0	0	0	0	0	17	0	17	0	6
	Insufficient	100	100	66	88	100	67	0	83	100	50	92	74
	Misconception	0	0	33	12	0	33	0	0	0	33	8	10

This decreased to 5% at the end year four. The reverse situation was observed in the claims produced among the second group of participants. While only 2% of the claims presented by the second group were wrong at the end of year two, this percentage rose up to 38% at the end of year four. The claims having the highest number of misconceptions were the second group discussing free fall. The claims having misconceptions decreased in general at the end of year four, and no claims were detected to have misconceptions in the fifth group's discussion.

When claims were analyzed in terms of being sufficient, insufficient, and having misconceptions, no warrants that addressed and supported the claim from many aspects were identified in any of the group discussions made at the end of year two. There were participants who supported their claims from only one aspect. Only in the fourth group's discussion was there sufficient warrants at the end of year four:

> *I begin first. Water, pure water boils at 100 degrees centigrade under atmospheric pressure. The characteristics of water, its purity, and its location and therefore its atmospheric pressure, they all change the boiling point. So, the differences in measurements may be due to these. The purity of water may be disrupted, changed or the atmospheric pressure where the boiling point is being measured may be increased or decreased. Thus, for the option A, I believe the different measurements may be the consequences of these changes.* (4Y-4E)

However, the rate of the warrants having misconceptions did not decrease in this group. It was found that the arguments examined were also of insufficient quality like the warrants. The fourth group developed arguments with no data at all at the end of year two. When looking at it from this viewpoint, the groups, however, produced a higher number of arguments using more data at the end of year four. Although misconceptions were not embedded in the data that much, misconceptions were identified to be present in the second group's discussion at the end of both years and in the third group at the end of year four. The data deemed to be sufficient were observed in the fourth and fifth groups' discussions at the end of year four.

Understanding the Nature of Scientific Argumentation: Competing Theories

When the arguments presented in the discussion on conflicting theories were examined in terms of quality, the rate of scientifically wrong claims decreased at the end of year four and, although some claims had misconceptions in the discussions conducted at the end of year two, there were no claims with misconceptions at the end of year four. The wrong claims were usually constructed when option F in the question was discussed:

> *I think it would deteriorate because the air flows through here, the s pipe, that is.* (2Y-1F)

At the end of year two, there was only one participant from this group who presented a moderate argument which included claim, data, and warrant; however, the claim made for the argument was not scientifically right and its warrant was weak:

> *There is a matter that serves as an insulating compound; therefore, as the soup will not contact with air, it can stay as it is forever.* (2Y-1D)

The warrants having misconceptions were present at the end of year two, but there were none detected at the end of year four. The warrants were usually based on the prevention of air contact. This warrant was presented on an activity sheet and the participants did not produce any other at the end of years two and four:

> *It can endure for a few days since it has no contact with air.* (4Y-1D)

The only different warrant presented was found to be the one that was produced at the end of year two, but with misconceptions:

> *The speed of putrefaction decreases related to working of enzymes.* (2Y-1A)

When the discussion process was analyzed, the other issue that caught attention was that none of the participants was able to realize that the first theory given for the question had misconceptions and the other theory was scientifically valid. Instead, all of their focus was centered upon matching the theories to their statements. Also, all the data used in this group were weak.

Constructing an Argument: Free Fall

When the arguments presented in the discussion on free fall were examined in terms of quality, the rate of scientifically right claims did change at the end of years two and four, the rate of wrong claims increased at the end of year four, and the rate of claims with misconceptions decreased. It was identified that the scientifically wrong claims were mostly present in options A and C of the question. However, the claim and its warrant were scientifically right, although the rebuttal presented for the claim had misconceptions:

> *Of course, this is only true if the volume is big, because then the frictional force would apply more, so while the speed of the brick would decrease, the coin would speed up.* (2Y-2C)

The two strong arguments, identified while the argument elements were being examined, were produced by the same participant. The warrants this participant presented for the arguments were for option A of the question and explained the claim from only one aspect, but had misconceptions:

> *The volume of brick is bigger. Of course, if it is a full brick. As the volume of the brick is bigger, more frictional force applies.* (2Y-2C)

When the warrants of the arguments were analyzed, no sufficient warrants were identified supporting the argument from different aspects. Although the rate of claims with misconceptions was high, the misconceptions were usually detected in the warrants. The misconceptions emerged in the warrants when the participants were explaining free fall with the help of using the concepts of volume, density, and weight:

> *As the density of air is close to that of the paper, the fall would slow down at this point.* (2Y-2D)

> *I threw the pencil. It fell due to frictional force. I think there is no frictional force in the air.* (4Y-2C)

It was concluded that in addition to the misconceptions on free fall, the participants also had serious misconceptions with regard to the aforementioned concepts. At the same time, the warrants were weak and supported the claims from only one aspect:

> *It falls. It affects the volume.* (2Y-2A)

The data used by the participants to construct their arguments were scientifically insufficient and some also had misconceptions:

> *I threw the pencil. It fell due to frictional force. I think there is no frictional force in the air.* (2Y-2A)

Understanding an Argument: Structure of Particulate Matter

When the arguments presented in the discussion on structure of particulate matter were examined in terms of quality, the majority of claims were scientifically right. The rate of claims deemed to be right increased at the end of year four, whereas the claims having misconceptions decreased in number at the end of year four.

As for the warrants, it was observed that while the rates of warrants with misconceptions and insufficient warrants were very close, the misconceptions detected in the warrants decreased at the end of year four as they decreased in the claims as well. The participants used warrants or rebuttals in a very similar structure. The warrants were expressed by stating a phrase such as, "due to particulate structure" and rebuttals such as, "if there was no particulate structure":

> *… but if they did not have particulate structure, they could not split.* (4Y-3F)

> *We are able to split it (paper) so it should have many particles.* (2Y-3C)

It was this group discussion where the participants produced the arguments by using the highest amount of data. At the end of year two, however, the data used was insufficient, as was the case at the end of year four. Misconceptions were present only when the concept of molecules was mentioned in the data:

> … *These particles should be flexible or should shrink.* (4Y-3F)

At the end of year two, the participants used insufficient data when providing their arguments, whereas at the end of year four, the data also had misconceptions. The data were in general insufficient at the end of both years.

Interpreting Empirical Data: Boiling Point of Water

When the arguments presented in the discussion on the boiling point of water were examined in terms of quality, the rate of scientifically correct claims was high at the end of both years. While the rate of claims deemed scientifically wrong was 11%, no claims were found to be present in this category at the end of year four. No significant change was observed in the rate of claims with misconceptions. It was found that the wrong arguments were produced when the participants proposed how they obtained the data indicating the boiling point as 105 and 108 degrees:

> *I think it is close to the sea level.* (2Y-4Y)

> *It can be above the sea level.* (4Y-4A)

When the warrants of the arguments were analyzed, they were found to be insufficient in general; however, at the end of year four, this group was the first to provide sufficient warrants that supported the claim from several aspects:

> *I begin first. Water, pure water boils at 100 degrees centigrade under atmospheric pressure. The characteristics of water, its purity, and its location and therefore its atmospheric pressure, they all change the boiling point. So, the differences in measurements may be due to these. The purity of water may be disrupted, changed or the atmospheric pressure where the boiling point is being measured may be increased or decreased. Thus, for option A, I believe the different measurements may be the consequences of these changes.* (4Y-4E)

While there were no warrants that had misconceptions at the end of year two, they were present at the end of year four. In this group, only the fourth year students used data and, except for one, all the others were insufficient.

Predicting and Explaining: Parallel Electric Circuit

When the arguments presented in the discussion on parallel electric circuits were examined in terms of quality, the group at the end of year two supported the highest number of wrong claims; however, this changed at the end of year four and 85% of all the claims were scientifically right.

When the warrants were analyzed, while no sufficient warrants were used at the end of years two and four, the rate of warrants having misconceptions present at the end of year two dropped down at the end of year four. After having examined the arguments comprising claims and warrants produced throughout the group discussion, it was concluded that the warrants were right yet they were insufficient and thus the participants could not provide full explanations and mostly used wrong claims:

> *They are both identical (bulbs) and the same current flows through both, if bulb A's connection is loosened, then more current will flow through bulb B.* (2Y-5A)

The same misconceptions were also found in the warrants and these misconceptions in turn caused the claim to be wrong:

As more current will flow through bulb B, it will be brighter. (2Y-5F)

These data used for the claims were insufficient. When the arguments were analyzed in terms of their scientific quality, it was observed that the participants' misconception was that they used the concepts of current and energy as the same. Only at the end of year four, use of sufficient data was identified.

Conclusion and Discussion

This study investigates the preservice elementary teachers' ability to use the components of an argument structure on given science topics as well as the quality of the developed argument contents at their sophomore and senior. In doing so, it was also aimed to discuss the effect of teacher education programs on the development of argumentation skills of students from years two and four. In this section, the results are discussed in the line of research questions.

The Components of Argument Structures

Toulmin's argumentation pattern was used for the determination of argument structure of the scientific discussions. According to this pattern, an argument's basic components are claim, warrant, and data. However, in more complex arguments, one can encounter qualifiers, backings, and rebuttals as well (Driver et al., 2000). As a result of the discussions analyzed in line with this model, it was found that the arguments proposed by the participants at the end of year two were in general composed of only claim or claim and warrants and were either very weak or weak. Whereas, at the end of year four, it was observed that the participants additionally used data in their arguments which involved claim and warrants and that these arguments were at moderate level. The participants attached particular importance to the use of claims when developing their arguments. When the study conducted by Jimenez-Aleixandre et al. (2000) with high school students was examined, it was seen that the results were similar to those obtained in this study. The students usually focused on the use of claims in their arguments on genetics and instead gave less weight to warrants or evidence to support the claim. Some preservice teachers who participated in the present study (i.e., 1E, 2D, 4B) provided only claims and they used no warrants or data at all. And yet, an argument that lacks warrants and is constructed only around a claim is not valid in general terms (Kaya ve Kilic, 2008). The other preservice teachers, however, did usually use warrants to support their claims. Proposal of numerous claims or claims and warrants was an indication that the participants did not have any problems in producing and supporting their own ideas. These results show similarities to some of the research conducted in the related literature (Aslan, 2014; Cinici et al., 2014; Demirci, 2008; Jimenez-Aleixandre & Rodriguez 2000; Kaya, 2012; Kutluca et al., 2014; Osborne et al., 2004). According to Demicioglu and Ucar (2014), the greater the quality of knowledge on a topic, fewer total number of claims are produced. In this study, too, the results were similar. While the rate of claims produced was as high as 40% at the end of year two, this rate decreased to 28% at the end of year four. Thus, it can be deducted that the content knowledge of those participants who only presented claims on the related subject was not as sufficient so as to help them develop quality and complex arguments (Aslan, 2014).

It is essential to make use of data and convincing evidence to support the validity of a claim when constructing a scientific argument (Clark & Sampson, 2008; Karisan & Yilmaz-Tuzun, 2012; Yildirir & Nakiboglu, 2014). Though increased, the research shows that the rate of arguments produced during all the discussions with data in addition to claims was still quite low (23%). This shows that the preservice teachers were unable to fully correlate claim and data which should be at the core of an argument, as they proposed warrants for their claims without using enough data or evidence. This is because the warrants are in fact the correlations established between the claim and the data (Kaya et al., 2014). In this case, we can argue that the students may actually have inadequate content knowledge on the subject at hand as they only made a claim or did not benefit from data and evidence to establish the claim-data relation.

When Demircioglu and Ucar (2014) examined the relationship between content knowledge and use of data, they found that the total amount of data used increased at the end of the posttests. The researchers arrived at the conclusion that the students were able to enhance their content knowledge throughout the study process and thus could create claims with more scientific reasoning and evidence. According to Sampson and Clark (2011), the students possessing a higher level of competencies in terms of scientific knowledge can foster the generation of more quality and complex arguments. This research, too, yielded similar results. In the arguments constructed at the end of year four, there were more data generated. When this is taken into account from the viewpoint of preservice elementary teachers, although they had completed the sciences courses at the end of year two, the science pedagogy helped them make up their inadequacies in content knowledge.

Another element to promote the quality of an argument for which warrant is made is rebuttals (Driver, et al., 2000; Kaya & Kilic, 2008; Osborne et al., 2004). The rate of arguments that included rebuttals among the statements deemed as arguments in the research was evenly low (11%) at the end of both years. Of the total number of arguments, only 2% of arguments had claims, warrants, and rebuttals and in only 1% of the arguments was data used in addition to said components. At the end of year four, this rate slightly changed, with the rate of arguments constructed with claims, warrants, and rebuttals being 4% and the arguments supported also with data being 1%. The rebuttals were usually presented as counter-claims. The highest number of rebuttals was produced in the second group. Demircioglu and Ucar (2014) claim that the increased number of rebuttals in the construction of arguments and that this number being higher than the warrants and data are an indication that these students are good at refuting ideas and thus the quality of arguments increases. Despite their conclusion, when the discussions of the second group at the end of years two and four were analyzed, it was determined that the rate of very weak and weak arguments was very high (42%) at the end of year two, and the rate of weak (27%) and moderate (47%) arguments was higher than the rate of arguments in other categories. Besides, no participant in the group was able to construct a strong argument at the end of year four. Therefore, the high number of arguments do not signify that the presented arguments are strong (Maloney & Simon, 2006).

The difference between the numbers of arguments among groups lies primarily in the structure of the questions. Yet, this difference can also present itself among groups discussing the same question (Maloney & Simon, 2006). According to Demircioglu and Ucar (2014) proposing more than one claim on a subject may arise due to some participants' lack of knowledge on that specific subject. When looking at the intergroup argument construction skills, each group's number and quality of arguments differ. In some of the groups, participants expressed solely their own opinions, not paying attention to the validity or inaccuracy of the opinions given by others and they did not provide rebuttals for counter-claims, all of which in turn reduced the quality of arguments developed during the discussion process. However, the fifth group discussing electrical circuit, for instance, where a smaller number of arguments was produced, half of the arguments were at moderate level in general at the end of year two. This result supports the findings of the research carried out by Maloney and Simon (2006).

From the aforementioned results, it can be said that the argument construction skills of the participants engaged in the research were weak. The participants' construction of very weak and weak arguments could be related to various reasons. One of them could be that they had never been given argumentation-based training. It has been shown by a great deal of research that the process of developing a valid and strong argument is not a spontaneous one, rather it is a skill that you acquire through education and practice (Hogan & Maglienti, 2001; Kuhn, 1991). Aslan (2014), too, notes in his study that one of the reasons for the failure of students in constructing arguments is that they never had any experience in argument construction and argumentation-based trainings. The findings of the research note that the quality of arguments created by the participants who attended an argumentation-based training increased (Demircioglu and Ucar, 2010; Karisan & Topcu, 2016; Karisan & Yilmaz-Tuzun, 2012; Kececi et al., 2011; Kingir et al., 2010; Yesildag et al., 2010).

Another reason could be the lack of knowledge of the participants on the discussion subjects. As the preservice teachers did not make use of scientific terminology whilst in discussion, this, in particular, can be an indication of their insufficient knowledge. In the related literature, studies support this idea (Aslan, 2014; Sampson & Clark, 2011; Tavares et al., 2010; vonAufscnaiter et al., 2008) and the idea that content knowledge does not directly affect the quality of an argument (Eskin & Bekiroglu, 2009; Khun, 1991; Kutluca et al., 2014; Perkins, et al., 1991). In their study conducted with physics preservice teachers, Hakyolu and Ogan-Bekiroglu (2011) assert that the quality of scientific knowledge does not directly link to overall quality of argument. In their research, Clark and Sampson (2008) are of the view that these findings should also be evaluated in terms of discussion frequency, presence of different opinions of the participants, and their skills of conveying their previous knowledge, rather than just focusing on sufficient scientific knowledge.

The Quality of Arguments

Another objective of the research was to determine the quality of the argument components. To this end, the components were examined in such categories as scientifically right, wrong, sufficient, insufficient, or having misconceptions. When the arguments were analyzed, it can be seen that while the participants were successful at the selection of a claim in general at the end of years two and four, they used warrants that supported their claims from a single aspect. Overall, the groups were able to create correct claims. However, some of these claims had misconceptions embedded in them. Among the groups, except for the fifth group at the end of year two, the presence of wrong claims was rare. When the warrants and data were examined, those that supported the claim from many aspects were not identified. In all the groups, the participants presented only one warrant whilst supporting their claims. These warrants were mostly devoid of scientific terminology. The data used was insufficient. This result was also similar to that in the study conducted by Aslan (2014) with high school students. Aslan (2014) asserts that this result is rooted in the inability of students to make correlations between other concepts as they lacked adequate content knowledge. Similar reasons were found in this study as well. For example, it was observed that the participants did not really include concepts apart from the ones already given in the activity sheet handed out during the discussions. The concepts they used had misconceptions in general and most of the time were not related to the subject in question.

The participants of the second group proposed the highest number of claims with misconceptions and this also applied for the warrants. And this had an impact on the scientific validity of their arguments. It was observed that the participants were unable to fully explain what they meant by the concepts of weight, mass, frictional force, and gravity. The same misconceptions were also detected in the study done by Kocakulah and Kenar Acil (2011) with eighth graders. The participants in the fifth group had the highest number of wrong claims as opposed to other groups. The reason as to why is assumed to be their inability to distinguish the differences between the concepts regarding electrical circuits. As the discussion ended within a very short period of time and the participants only presented their opinions and did not further discuss them, there was not much data. The concepts such as current, voltage, and energy identified in the warrants were used as if they bear the same meaning. These results show similarities with that of Kucukkozer's (2003) research administered to high school students.

The Contribution of The Teacher Education Programs

In Turkey, teacher training program courses did not directly support argumentation-based education. On the other hand, preservice teachers are required to engage in scientific discussions as a part of their courses to enrich their knowledge in science and the teaching of science. Since the preservice teachers participated in the science courses for the first two years and that they followed up the Science Laboratory Practices I and II during the academic year when the study was under way, it was assumed that their content knowledge would be enough to address the discussion subjects at primary school level. Whereas Means and Voss (1996) argue that sufficient scientific knowledge does not influence all the aspects of an argument as much as it affects the backings in particular. As for this study, although the content knowledge which the preservice teachers had was sufficient at the end of year two, it was also deducted that the enhancement of their arguments' quality at the end of year four showed that they were better at comprehending and transferring this knowledge after receiving technology and pedagogy courses where they used their knowledge in practice. The practices embedded in these courses enabled the preservice teachers to present many scientific concepts they had learned previously in theory. It can be argued that these preparations helped them have command of subjects within science teaching.

However, the difficulties they encounter in forming arguments can be attributed to their not having the opportunity to form enough arguments in the courses they had taken during their education program. In the research conducted by Aydemir et al. (2018) with science teachers, they found that preservice science teachers also formed insufficient arguments. The preservice teachers stated that traditional teaching was used intensively in undergraduate courses, the subjects were generally explained verbally, and the argumentation approach was used very rarely in the courses; therefore, they found themselves insufficient prepared in the argumentation approach and argumentation. In support of this view, Sahin et al. (2015) stated in their study with faculty of education that the academics who run science classes carry the argumentation method to their classes, but they often have difficulty due to the crowded classes and the content of the course.

The science laboratory I and II courses, which have a great potential for creating arguments, are designed as a course in which teacher candidates do their laboratory applications. The studies related to these courses reveal that argumentation and inquiry-based learning practice processes contribute to laboratory applications by helping in structuring scientific knowledge (Hohenshell & Hand, 2006). However, research reveals that preservice elementary teachers do not frequently encounter argumentation-based practices during their education (Cappellaro, 2016). Cappellaro (2016) also found that half of the participants who completed the Physics, Chemistry, Biology, Science Laboratories I and II, and Science Teaching I courses stated that they did not take a course that used argumentation-based teaching. Preservice teachers also think that a scientific classroom environment can be created using direct instruction and experimentation. They also mentioned the laboratory as the primary method that can be used in science where argumentation was mentioned only by one of the 36 preservice teachers who participated in the study (Karaer et al., 2020). Among the reasons why these preservice teachers who participated in this study could not produce persuasive and qualified arguments in the discussions in real terms was that they had not received any training aimed at developing their argumentation skills and therefore failed to fully participate in such discussion activity. The second reason could be that the learning settings where preservice teachers can transfer and question the knowledge they acquired, where their awareness can be raised to realize the misconceptions that they or their students' arguments may have, and where scientific issues are addressed and discussed are not provided as much as they should be. From this aspect, it can be concluded that providing preservice teachers with settings where they can produce more arguments during the courses they take and where they can engage in discussions would build their capacity for constructing arguments and help them perceive their misconceptions (Kingir et al., 2011).

Considering the critical role of elementary teachers in preparing students to think, explain, ask, and argue about science and interconnected concepts, science teachers at elementary level need more specific attention considering the variety of courses they must teach. This study contributes to an understanding of preservice elementary teachers' argumentation skills and the role of teacher education programs. According to the results, developing preservice teachers' understanding and experiences in teaching science through argumentation is necessary and intervention programs as a part of science courses also need to be considered. Further research needs to deal with instructional and practical issues.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Evren Cappellaro (evrenne@gmail.com) is a lecturer in the department of elementary education at Akdeniz University where she teaches science education and environment education courses to teacher candidates. Her research focuses on scientific thinking, the nature of science, and environmental issues with a particular focus on water-related environmental science.

Demet Seban (demet.seban@alanya.edu.tr) is an associate professor in the Department of Elementary Education at Alanya Alaaddin Keykubat University. Her interest areas include curriculum and instruction, teacher education, and teaching literacy. More specifically, she works on preparing teacher candidates to teach reading and writing, content area literacies, and developing language education programs at the elementary level.

References

Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S., Reznitskaya, A., Tillmanns, M., & Gilbert, L. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction, 19*(1), 1–46.

Aslan, S. (2014). Ogrencilerin yazili bilimsel arguman olusturma ve degerlendirme becerilerinin incelenmesi [Analysis of students' written scientific argument generate and evaluation skills]. *Journal of Theory and Practice in Education, 10*(1), 41-74.

Aydemir, S., Karakaya Cirit, D., Kaya, S., & Azger, C. (2018). Fen bilgisi ogretmen adaylarinin argumantasyona iliskin gorusleri ve argüman kurma becerilerinin arastirilmasi [Exploring the preservice science teachers' views of argumentation and of their argument development skills]. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi, 6*(STEMES'18), 131-138. https://doi.org/10.18506/anemon.470577

Cappellaro, E. (2016, May). *Sinif ogretmeni adaylarinin argumantasyon algıları* [Preservice elementary teachers' argumentation perceptions. [Paper presentation]. 15th International Primary Teacher Education Symposium, Mugla.

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

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*, 427–449. https://doi.org/10.1080/09500690802627277

Cinici, A., Ozden, M., Akgun, A., Herdem, K., Karabiber, H. L., & Deniz, S. M. (2014). Kavram karikaturleriyle desteklenmis argumantasyon temelli uygulamaların etkililiginin incelenmesi [Investigating the effectiveness of argumentation-based activities supported with concept cartoon]. *Adiyaman Universitesi Sosyal Bilimler Enstitusu Dergisi, 7*(18), 571- 596. https://doi.org/10.14520/adyusbd.839

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*, 293-321. https://doi.org/10.1002/tea.20216

Cresswell, J. W. (2007). *Qualitative inquiry and research design* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Demirci, N. (2008). *Toulmin'in bilimsel tartısma modeli odakli egitimin kimya ogretmen adaylarinin temel kimya konularini anlamalari ve tartisma seviyeleri uzerine etkisi* [The impact of Toulmin's focus on the scientific debate model of chemistry teacher candidates on understanding and discussing basic chemistry subjects]. (Unpublished master's thesis). Gazi Universitesi Egitim Bilimleri Enstitusu, Ankara.

Demircioglu, T., & Ucar, S. (2014). Akkuyu Nukleer Santrali konusunda uretilen yazili argumanlarin incelenmesi [Investigation of written arguments about Akkuyu Nuclear Power Plant]. *Elementary Education Online, 13*(4), 1373-1386. https://doi.org/10.17051/io.2014.31390

Denzin, N. K., & Lincoln, Y. S. (1998). Entering the field of qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research theories and issues* (pp. 1-34). Thousand Oaks, CA: Sage Publications.

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

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education, 38*, 39-72. https://doi.org/10.1080/03057260208560187

Erduran, S., Ozdem, 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. https://doi.org/10.1186/s40594-015-0020-1

Eskin, H., & Bekiroglu, F. O. (2009). Investigation of a pattern between students' engagement in argumentation and their science content knowledge: A case study. *Eurasia Journal of Mathematics, Science and Technology Education, 5*(1), 63-70.

Hakyolu, H., & Ogan-Bekiroglu, F. (2011). Assessment of students' science knowledge levels and their involvement with argumentation. *International Journal for Cross- Disciplinary Subjects in Education, 2*(1), 264-270. https://doi.org/10.20533/ijcdse.2042.6364.2011.0038

Hohenshell, l. M., & Hand, B. (2006). Writing-to-learn strategies in secondary school cell biology: A mixed method study. *International Journal of Science Education, 28*(2-3), 261-289.

Hogan, K., & Maglienti, M. (2001). Comparing the epistemological underpinning of students' and scientists' reasoning about conclusions. *Journal of Research in Science Teaching, 38*(6), 663-687. https://doi.org/10.1002/tea.1025

Jimenez-Aleixandre, M. P., Rodriguez, A. B., & Duschl, R. A. (2000). "Doing the lesson" or "doing science": Argument in high school genetics. *Science Education, 84*, 757-792. https://doi.org/10.1002/1098-237X(200011)84:6<757::AID-SCE5>3.0.CO;2-F

Jimenez-Aleixandre M. P., & Erduran, S. (2008). Argumentation in science education: An overview. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education perspectives from classroom-based research*, (pp. 3–28), Netherlands: Springer.

Karaer, G., Karademir, E. ve Tezel, O. (2020). Proje tabanli ogretimin fen laboratuvari dersinde kullaniminin sınıf ögretmeni adaylarının goruslerine gore degerlendirilmesi [Evaluating the use of project based teaching on the science laboratory course according to the views of the classroom teacher candidates]. *Journal of Qualitative Research in Education*, *8*(1), 73-95. https://doi.org/10.14689/issn.21482624.1.8c.1s.4m

Karasar, N. (1995). *Bilimsel arastırma yontemi* [Scientific research method]. Ankara: 3A Arastirma Egitim Danismanlik Ltd. Sti.

Karisan, D., & Topcu, M. S. (2016). Contents exploring the preservice science teachers' written argumentation skills: The global climate change issue. *International Journal of Environmental and Science Education, 11(6),* 1347-1363. https://doi.org/10.12973/ijese.2016

Karisan, D., & Yilmaz-Tuzun, O. (2012, June 27). *Fen ogretmen adaylarinin laboratuvar uygulamalari dersi deney raporlarinin incelenmesi* [Examination of experiment reports of laboratory practice lessons of science teacher candidates]. [Paper presentation]. X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde.

Kaya, E. (2012, June 27). *Argumantasyona dayali etkinliklerin ogretmen adaylarının kimyasal denge konusunu anlamalarına etkisi* [The influence of the argumentation-based activities on the students' understanding of the chemical equilibrium]. [Paper presentation]. X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde.

Kaya, E., Cetin, P. S., & Erduran, S. (2014). Iki argumantasyon testinin Turkce'ye uyarlanmasi [Adaptation of two argumentation tests into Turkish]. *Elementary Education Online, 13*(3), 1014-1032.

Kaya, O. N., & Kilic, Z. (2008). Etkin bir fen ogretimi için tartısmaci soylev [Argumentative discourse for the effective teaching of science]. *Ahi Evran Universitesi Kirsehir Egitim Fakultesi Dergisi, 9*(3), 89-100.

Kececi, G., Kirilmazkaya, G., & Kirbag Zengin, F. (2011, May 16). *Ilkogretim ogrencilerinin genetigi degistirilmis organizmalari on-line argumantasyon yontemi ile ogrenmesi.* [Elementary school students' learning about genetically modified organisms with the online scientific argumentation learning program]. [Paper presentation]. 6th International Advanced Technologies, Elazig.

Kingir S., Geban, O., & Gunel M. (2011). Ogrencilerin kimya derslerinde argumantasyon tabanli bilim ogrenme yaklasiminin uygulanmasina iliskin gorusleri [Students' views on the implementation of the argument-based science learning approach in chemistry classes]. *Selcuk Universitesi Ahmet Kelesoglu Egitim Fakultesi Dergisi, 32*, 15-28.

Kingir, S., Geban, O. & Gunel, M. (2010, September 23). *Argumantasyon tabanli bilim ogrenme (ATBO) yaklasiminin 9. sinif ogrencilerinin kimya kavramlarini ogrenmelerine etkisi* [The impact of the argument-based science learning (ABSL) approach to learning chemistry concepts of 9th grade students]. [Paper presentation]. IX. Ulusal Fen Bilimleri ve Matematik Kongresi, Izmir.

Kim, S., & Hand, B. (2015). An analysis of argumentation discourse patterns in elementary teachers' science classroom discussions. *Journal of Science Teacher Education, 26*(3), 221-236. https://doi.org/10.1007/s10972-014-9416-x

Kocakulah, S., & Kenar Acil, Z. (2011). Ilkogretim ogrencilerinin gozuyle "Yercekimi neredеdir?" [From the perception of elementary school students: "Where is gravity?"]. *Turk Fen Egitimi Dergisi, 8*(2), 135-151.

Kucukozer, H. (2003). Lise I. sinif ogrencilerinin basit elektrik devreleriyle ilgili olarak kavram yanilgilari [Misconceptions of first year secondary school students' about simple electric circuits]. *Hacettepe Universitesi Egitim Fakultesi Dergisi, 25*, 142-148.

Kuhn, D. (1991). *The skills of argument.* England: Cambridge University Press.

Kuhn, D. (1992). Thinking as argument. *Harvard Educational Review*, 62, 155-178.

Kuhn, D. (2005). *Education for thinking.* Cambridge, MA: Harvard University Press.

Kutluca, A. Y., Cetin, P. S., & Dogan, N. (2014). Effect of content knowledge on scientific argumentation quality: Cloning context. *Necatibey Egitim Fakultesi Elektronik Fen ve Matematik Egitimi Dergisi, 8*(1), 1-30. https://doi.org/10.12973/nefmed.2014.8.1.a1

Lee, M-H., Wu, Y., Tien, T., & Chin-Chung, A. (2009). Research trends in science education from 2003 to 2007: A content analysis of publications in selected journals. *International Journal of Science Education, 31*(15), 1999–2020. https://doi.org/10.1080/09500690802314876

Lin, T. C., Lin, T. J., & Tsai, C. C. (2014). Research trends in science education from 2008 to 2012: A systematic content analysis of publications in selected journals. *International Journal of Science Education, 36*(8), 1346–1372. https://doi.org/10.1080/09500693.2013.864428

Maloney, J., & Simon, S. (2006). Mapping children's discussion of evidence in science to assess collaboration and argumentation. *Science Education, 28*(15), 1817-1841. https://doi.org/10.1080/09500690600855419

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203-229. https://doi.org/10.1002/sce.20364

Means, L. M., & Voss, J. F. (1996). Who reasons well? Two studies of informal reasoning among children of different grade, ability and knowledge levels. *Cognition and Instruction, 14*(2), 139-178.

M.E.B. (2015). Fen bilimleri dersi (3-8. sınıflar) öğretim program [Science Curriculum (3-8. Grades)]. Ankara: Milli Egitim Bakanligi.

Okada, A., & Buckingham Shum, S. (2008). Evidence-based dialogue maps as a research tool to evaluate the quality of school pupils' scientific argumentation. *International Journal of Research & Method in Education, 31*(3), 291–315. https://doi.org/10.1080/17437270802417184

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

Osborne, J. F., Erduran, S., Simon, S., & Monk, M. (2001). Enhancing the quality of argument in school science. *School Science Review, 82*(301), 63-70.

Ozdem, Y., Cakiroglu, J., Ertepinar, H., & Erduran, S. (2013). The nature of preservice science teachers' argumentation in inquiry-oriented laboratory context. *International Journal of Science Education, 35*(15), 2559–2586. https://doi.org/10.1080/09500693.2011.611835

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In Voss, J. F., Perkins, D. N., & Segal, J. W. (Eds.), *Informal reasoning and education*, (pp. 83-105). Hillsdale: Erlbaum.

Plantin, C. (2005). *L'Argumentation. Histoire, théories et perspectives.* Paris: Presses Universitaires de France.

Sahin, C., Cakmak, N., & Hacimustafaoglu, M. (2015). Akademisyenlerin elestirel dusunme becerilerinin ve elestirel dusunme becerileri hakkindaki goruslerinin incelenmesi [Investigation of academicians' critical thinking skills and ideas about critical thinking skills]. *Inonu Universitesi Egitim Fakultesi Dergisi, 16*(2), 51-66. https://doi.org/10.17679/iuefd.16211687

Sampson, V., & Clark, D. (2011). A comparison of the collaborative scientific argumentation practices of two high and two low performing groups. *Research in Science Education, 41*, 63-97. https://doi.org/10.1007/s11165-009-9146-9

Sampson, V., & Clark, D. (2008). Assessment of the ways students generate arguments in science education: Current perspectives and recommendations for future directions. *Science Education, 92*(3), 447-472. https://doi.org/10.1002/sce.20276

Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *The Journal of the Learning Sciences, 12*(2), 219-256. https://doi.org/10.1207/S15327809JLS1202_3

Tavares, M. L., Jimenez-Aleixandre, M. P., & Mortimer, E. F. (2010). Articulation of conceptual knowledge and argumentation practices by high school students in evolution problems. *Science and Education, 19*(6), 573-598. https://doi.org/10.1007/s11191-009-9206-6

Toulmin, S. (1958). *The uses of argument.* Cambridge, UK: Cambridge University Press.

Tumay, H., & Koseoglu, F. (2011). Kimya ogretmen adaylarinin argumantasyon odakli ogretim konusunda anlayislarinin gelistirilmesi [Improving the understanding of chemistry teacher candidates on argumentation-based teaching]. *Turk Fen Egitimi Dergisi, 8*(3), 105-119.

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. https://doi.org/10.1002/tea.20213

Walton, D. N. (1996). Argumentation Schemes for Presumptive Reasoning. Mahwah, NJ: Erlbaum.

Walton, D. N., Reed, C., & Macagno, F. (2008) Argumentation Schemes. Cambridge: Cambridge University Press.

Yesildag, F., Gunel, M., Yilmaz, A. (2010, September 23). *Ilkogretim 8. sinif seviyesinde maddenin yapisi ve ozellikleri unitesini ogrenmede argumantasyon tabanli bilim ogrenme (ATBO) yaklasiminin akademik basariya etkis*i [The impact of argumentation based science learning (ABSL) approach on academic achievement in learning the unit of substances and their properties at 8th grade]. [Paper presentation]. IX. Ulusal Fen Bilimleri ve Matematik Kongresi, Izmir.

Yildirir, H. E., & Nakiboglu, C. (2014). Kimya ogretmen ve ogretmen adaylarının derslerinde kullandiklari argumantasyon sureclerinin incelenmesi [Examination of chemistry teachers and preservice teachers' argumentation processes used in their courses]. *Abant Izzet Baysal Universitesi Egitim Fakultesi Dergisi, 14*(2), 124-154. https://doi.org/10.17240/aibuefd.2014.14.2-5000091531