ISSN 1303-0485 • **eISSN** 2148-7561
DOI 10.12738/estp.2015.1.2324
Copyright © 2015 EDAM • http://www.estp.com.tr
Educational Sciences: Theory & Practice • 2015 February • 15(1) • 267-283

Received | 6 December 2013
Accepted | 18 September 2014
OnlineFirst | 15 January 2015

Investigating the Effect of Argument-Driven Inquiry in Laboratory Instruction*

Tuba Demircioglu[a]
Cukurova University

Sedat Ucar[b]
Cukurova University

Abstract

The aim of this study is to investigate the effect of argument-driven inquiry (ADI) based laboratory instruction on the academic achievement, argumentativeness, science process skills, and argumentation levels of pre-service science teachers in the General Physics Laboratory III class. The study was conducted with 79 pre-service science teachers. The participants in the control group (n = 38) participated in traditional laboratory activities, and the participants in the experimental group (n = 41) participated in laboratory activities based on argument-driven inquiry. Data was collected through the Optical Achievement Test (OAT), Argumentative Scale (AS), Science Process Skills Test (SPST) and the individual reports of the participants. Qualitative and quantitative techniques were used together to analyze the data. The results showed that argument-driven inquiry was more effective in improving the academic achievement and science process skills of pre-service science teachers compared to traditional laboratory instruction, but no significant difference was observed in the Argumentative Scale scores between the groups that had ADI instruction and those that had traditional laboratory instruction. Towards the end of the treatment, there was an improvement in the argumentative quality of the experimental group, but there was no change in the argumentation quality of the control group. ADI is an effective method for improving the academic achievement and science process skills of students, and it could be adapted for other laboratory classes. Argumentativeness might be improved with a longer argumentation session and more laboratory activities.

Keywords: Argumentation • Argument-driven inquiry • Laboratory instruction • Teacher education • Geometrical optics

---

* This study is a part of Tuba Demircioglu's master thesis supported by Cukurova University Scientific Research Projects with grant number of EF2010YL20.

a **Corresponding author**
Tuba Demircioglu, Cukurova University, Faculty of Education, Department of Elementary Education, 01330 Saricam, Adana, Turkey
Research areas: Science education, teacher education, argumentation
Email: tdemircioglu@cu.edu.tr

b Sedat Ucar, Cukurova University, Faculty of Education, Department of Elementary Education, 01330 Saricam, Adana, Turkey
Email: sucar@cu.edu.tr

After the Soviet Union launched Sputnik in 1957, there was an inquiry about the reasons for the technological gap between the United States and the Soviet Union in the space race. There appeared a need for reform in education, especially in science and mathematics. This led to a revision in educational programs from primary school to institutions of higher education in the United States (Hiatt, 1986). With the post-sputnik era, scientific inquiry and laboratory training became an important part of science education (Anderson, 2007; Hiatt, 1986).

Scientific inquiry is the basis for research and study (Anderson, 2007; Cobern et al., 2010), and argumentation is one of the most important processes of scientific inquiry (Sampson, Grooms, & Walker, 2011). The evaluation and interpretation of evidence, the evaluation of the validity of scientific knowledge, and thinking about different ideas are the core elements of argumentation and science. They play an important role in the construction of scientific knowledge (Diver, Newton, & Osborne, 2000; Duschl & Osborne, 2002). In science education, students should engage in activities that require them to use effective language and perform scientific reasoning with their peers and teachers. This means participating in the construction and evaluation of scientific argumentation (Duschl & Osborne, 2002).

Argumentation plays a crucial role in the construction of scientific explanation and creation of theories. Scientists engage in argumentation to create and improve scientific knowledge (Aufschnaiter, Erduran, Osborne, & Simon, 2008; Nussbaum & Sinatra, 2003). Engaging in the process of argumentation requires students to make claims, use data to support their claims, and justify claims with scientific evidence. With this process, students learn science concepts and have the opportunity to practice the methods used by scientists to justify or refute their claims. During scientific argumentation, students reflect their own ideas and learn about the ideas of others. Hence, it helps to correct misconceptions and ensures a meaningful learning experience (Aufschnaiter et al., 2008).

Scientific inquiry and argumentation has a crucial place in science education (Duschl & Osborne, 2002; Zohar & Nemet, 2002). Argumentation is a discussion format that needs to be taken seriously by students and be taught explicitly in science classes through appropriate teaching and modeling methods (Duschl & Osborne, 2002; Jiménez-Aleixandre & Erduran, 2008; Jiménez-Aleixandre, Rodriguez, & Duschl, 2000; Kelly, Druker, & Chen, 1998). Although argumentation has an important place in science education, it is rarely used in science courses or laboratory activities (Driver et al., 2000; Jimenez-Alexander et al., 2000; Newton, Driver, & Osborne, 1999; Wellington & Osborne, 2001).

In recent years, a great number of studies have been carried out on the implementation of argumentation in science classes (Aufschnaiter et al., 2008; Driver et al., 2000; Duschl & Osborne, 2002; Kelly & Takao, 2002; Zohar & Nemet, 2002). In those studies, it was highlighted that the reasoning skills, argument and counter-argument construction skills of teachers and pre-service science teachers, as well as teacher strategies to engage students in argumentation were faulty (Driver et al., 2000; Zeidler, 1997; Zohar, 2008). Lots of science teachers have problems integrating argumentation and using scientific inquiry in their class, as well as engaging students in scientific inquiry to help them understand the development of important concepts in science (Sampson & Gleim, 2009; Simon, Erduran, & Osborne, 2006). Studies show inconsistent results in the development of argumentation skills of students who engage in scientific inquiry activities. As a result of the analysis carried out after the laboratory practices of Kelly et al. (1998) on electrical circuits with mystery boxes, it was found that most of the time students completed their arguments without warranting, and in this respect, these are fallacies. Failures in warranting the arguments or partial warranting are also attributed to a lack of knowledge. On the other hand, in the study of Kim and Song (2005) in which they investigated the arguments of students engaged in scientific inquiry activities, it was observed that students had improved their argumentation process. Watson, Swain, and McRobbie (2004) found results contrary to Kim and Song. In the study of Watson et al., the quality and quantity of the arguments of scientific inquiry were low. It was believed that this study could help individuals with difficulties in scientific inquiry, argumentation, and their use in the classroom by providing information about scientific inquiry and argumentation. Although science laboratory practices aim to provide concrete experiences, activate experiential learning, increase student knowledge, and provide meaningful and permanent learning, the "cookbook" approach in traditional laboratory practices hinders science classes from reaching their goals (Hoffstein & Lunetta, 2004; Schen, 2007; Walker, Sampson, Grooms, Anderson, & Zimmerman, 2010). Similarly, the traditional deductive laboratory approach which explicitly gives information on how to carry out an experiment, collect and analyze the data, and indirectly tells what results should be found in relation to the theoretical knowledge is also widely used in Turkey (Aydoğdu & Ergin, 2008; Feyzioğlu et al., 2011; Yağbasan & Kanlı, 2008). However, instead of "cookbook" laboratory activities in which each step is given by the lab manual, students need laboratory activities in which they can inquire, suggest and test hypotheses, share their ideas clearly, and do scientific research like real scientists (Demircioğlu, & Ucar, 2012). In this regard, it is clear that the success of laboratory instruction is closely related with the techniques used. Hence, various techniques have been used and their effectiveness has been evaluated (Lunetta, Hoffstein, & Clough, 2007). It has been found that in studies using Science Writing Heuristic (SWH), a mehod which combines research and inquiry techniques with the basic language learning processes like reading writing and speaking, the academic achievement of students increases by ensuring the production of scientific knowledge through argumentation (Akkuş, Günel, & Hand, 2007; Burke, Hand, Poack, & Greenbowe, 2005). Cross, Taasoobshirazi, Hendricks, and Hickey (2008) stated that argumentation and collaborative group studies affect learning and success in science. Zohar and Nemet (2002) confirmed that the experimental group students engaged in argumentation increased their conceptual knowledge test scores.

One of the methods suggested for increasing student achievement in laboratory instruction is Argument-Driven Inquiry (ADI). ADI is a method similar to models such as SWH and the 5E learning cycle models, and provides students with the opportunity to construct their own explanations and share their ideas while socializing in small groups or during in-class discussions. It creates a classroom atmosphere that provides a cultural process for the teaching of science (Sampson, 2009). This model is designed to make laboratory instruction more informative and to plan scientific inquiry which includes argument development through research questions (Sampson, 2009; Sampson & Gleim, 2009; Walker et al., 2010). This method gives students an opportunity to generate an argument in the direction of their own questions, to assert methods for finding answers to those questions, and to design the research through group work. It provides students with an opportunity to evaluate the validity and reliability of their data, and to spend time on key concepts and ideas in order to form a deep understanding (Sampson, 2009; Sampson & Gleim, 2009; Walker et al., 2010). ADI is different from other methods in that it provides students with a chance to design their research and find results on their own, as well as to engage in the argumentation process where they can share and support their ideas. This method consists of peer reviews which improve the critical thinking skills of students; it provides students with the opportunity to see and fix their shortcomings. ADI is additionally thought to be an effective method since it provides students with the ability to share and evaluate their products with each other, improve their communication and writing skills, understand the construction of scientific knowledge, and have a chance to experience things in person. In laboratory classes which use ADI, it has been found that students are able to use scientific evidence better, they have better reasoning skills and a more positive attitude towards science (Sampson, Walker, Dial, & Swanson, 2010), their peer review performance is good (Sampson et al., 2010), and they produce higher quality written arguments (Sampson et al., 2011). Unfortunately, there are few studies on Argument-Driven Inquiry in the literature. The main concern of these studies has been to focus on the argumentation skills of students, their attitudes towards science, and their writing skills. Thus, it is necessary to question the effectiveness of Argument-Driven Inquiry in regard to certain variables. In this study the effect of the ADI method on academic achievement, argumentativeness, scientific process skills, and the argumentation levels of students has been investigated.

Meaningful learning in science requires testing the usefulness of ideas needed for estimation, providing evidence to test the estimation or to find answers to the questions, and it requires interpretation of the results. In other words, it necessitates the use of scientific process skills (Harlen, 1999). For example, if scientific process skills are not improved during science education, then related evidence cannot be provided nor can results be constructed on findings which approve previous suppositions and reject counter-evidence. The new concepts might not be enough to understand the world. Hence, improving science process skills should be one of the main objectives of science education (Harlen, 1999). There is a relationship between experimental teaching strategies and improving science process skills (Marbie & Baker, 1996). While traditional laboratory methods give importance to outcomes, principles and learning the Truth, skills like problem solving, scientific thinking and inquiry are left out (Welch, Klopfer, Aikenhead, & Robinson, 1981). These kinds of experiments do not help students gain scientific process skills or construct scientific knowledge (Aktamış, 2007). According to Myers and Dyer (2006), students who take inquiry laboratory instruction got higher scores on science process skills when compared to those who took traditional laboratory instruction. Tümay and Köseoğlu (2011) designated that after an argumentation-focused chemistry class, pre-service science teachers thought argumentation can improve various skills such as inquiry and scientific thinking. Duru, Demir, Önen, and Benzer (2011) determined that inquiry-based laboratory practices show a positive increase in the skills of students who use science processes. Günel, Kıngır, and Geban (2012) stated that with the Science Writing Heuristic (SWH) approach, students construct knowledge in a research-inquiry based learning environment where they ask questions, build scientific claims, and support them with evidence. They claim that this approach forms a more active learning environment by providing students to participate in the learning process. It is important to investigate the contributions of studies in science teaching regarding the development of science process skills in laboratories that use the ADI method.

In studies, it can be seen that the structure of student arguments is weak and some students do not participate in argumentation (Jimenez-Alexandre et al., 2000; Kelly et al., 1998; Watson et al., 2004; Zohar & Nemet, 2002). There are many reasons that affect student participation in argumentation. Their *tendency to argue about controversial issues (argumentativeness)* might be one of them. Because individuals with a low tendency to argue about controversial issues avoid generating an argument, they are not comfortable engaging in argumentation (Infante & Rancer, 1982). In studies investigating argumentativeness, it is seen that the tendency of participants to argue about controversial issues increased significantly after becoming involved in argumentation practices (Kaya & Kılıç, 2008; Rancer, Whitecap, Kosberg, & Avtgis, 1997). The cross-sectional study of Schullery and Schullery (2003) analyzed the relation between age and educational background of the participants with argumentativeness, and it confirmed that as one gets older, their tendency towards argumentation decreases. If the argumentativeness of students is known, their inclusion in the argumentation process becomes possible. In this case, one subject that needs to be researched is the argumentativeness of students. The overall aim of this study is to investigate the effect of argument-driven inquiry (ADI) on the academic achievement, argumentativeness, scientific process skills, and level of argumentation of pre-service science teachers during the course *General Physics Laboratory III*.

Method

Research Design

In this research, the non-equivalent groups design from the quasi-experimental design was used. Unlike experimental designs, non-equivalent group members are not randomly chosen. They are, however, randomly assigned to the control or experimental groups (Cohen, Manion, & Morrison, 2007).

Study Group

The study was carried out during the fall semester of the 2010-2011 academic year. The study group consisted of students majoring in elementary school science teaching at a research university in southern Turkey. There were a total of 79 pre-service science teachers who participated in the study. 41 were in the experimental group, 26 females and 15 males. 38 were in the control group, 30 females and 8 males. Among the non-random sampling methods, convenience sampling was used in the study. Considering the fact that the researchers worked at that university, to eliminate time and work force limitations, the study was carried out using students from this university.

Data Collection Tools

Qualitative and quantitative data collection tools were used in the study. Quantitative data was collected before and after the training as pre and post-tests through the Optic Achievement Test (OAT), developed by the researcher, as well as the Argumentative Scale (AS), and Science Process Skills Test (SPST). On the other hand, qualitative data was collected through document analysis. Document analysis involved the analysis of the writings on information about the phenomenon being researched. In this regard, after each experiment that the students carried out, data was collected via individually written reports from the first, middle, and last practice. Later, these reports were examined and analyzed by the researchers using descriptive analysis.

Optical Achievement Test (OAT): To measure the geometrical optics achievement of sophomore pre-service science teachers in laboratory classes, the Optical Achievement Test was assigned as a pre-test and a post-test. The OAT was designed by the researcher and developed in the following 4 phases:

I. Experiments that would be carried out in the course of the General Physics Laboratory III class and the subjects that would be covered in this class under "Geometrical optics" were identified. Later, a test consisting of 36 multiple-choice questions related to those subjects was designed. The test included questions with five answer options.

II. Two associates from the Department of Elementary Science Education examined the content validity of the test. Necessary changes were applied in line with the views of experts.

III. At the beginning of the 2010-2011 fall semester, the test was given to 69 pre-service science teachers who had previously taken this course. In scoring the test, each correct item was scored as 1, and each incorrect or unanswered item was scored as 0.

IV. Item analysis using ITEMAN software was carried out for the pilot test. After the necessary changes, the final version of the test was readied and included 25 items. With a reliability value of .70 using the Kuder Richardson-20 (KR-20) scale, the OAT was considered reliable.

Argumentative Scale (AS): To measure teacher candidates' argumentativeness, the Argumentative Scale (AS), developed by Infante and Rancer (1982), was used. The five-point Likert-type questionnaire consisted of 10 items for the tendency to approach arguments (ARG_{ap}) (e.g. I am happy to defend my opinion on a topic), and 10 items for the tendency to avoid argument (ARG_{av}) (e.g. When I realize I am about to enter into a discussion I don't feel good) for a total of 20 items. Cronbach's alpha reliability for the Turkish adaptation of the scale by Kaya and Kılıç (2008) was .73. For this study, Cronbach's alpha reliability coefficient of the survey was .90 for the pretest, and .87 for the post-test. Therefore, the survey was evaluated as having a good level of reliability.

Science Process Skills Test (SPST): To measure the scientific process skills of pre-service science teachers, the Scientific Process Skills Test (SPST) developed by Okey, Wise, and Burns (1982) was used. The multiple choice test consisted of 36 questions with four options for each. These questions measured the following abilities: 12 questions for identifying variables in a problem, 8 for hypothesizing and defining, 6 for bringing procedural descriptions, 3 for designing the necessary investigations to solve the problems, and 7 questions for graphing and interpretation. The reliability of the test was .82. The test was translated and adopted into Turkish by Geban, Aşkar, and Özkan (1992). The reliability of the Turkish version of the test was .81 (Geban et al., 1992). Therefore, assessed as having a good level of reliability, the test was applied both to the experimental and control groups as a pretest and a post-test.

Student Reports: The experimental and control group participants prepared individual reports according to their data as well as their results from each experiment. In order to determine the pre-service science teachers' level of argumentation, these reports were examined. Reports of the experimental group participants consisted of three sections for answering the questions: "What were you trying to do and why?" "What did you do and why?" and "What is your argument?" The reports from the control group consisted of the following sections: Name of the Experiment, Purpose of the Experiment, Experimental Procedure, Data, Calculations, and Results and Comments.

Treatment

Experimental Group: Argument-Driven Inquiry Based Laboratory Method: Before the ADI laboratory course, the researcher divided participants into heterogeneous groups of four people (one group had five people) in terms of gender and success. Prior to each laboratory course, participants were regrouped with different peers to force all participants to work in different groups each time. Participants carried out the activities in groups for the first six stages of the method. All participants in the class were divided into two groups. In one group for the first two hours and the other group for the second two hours, they carried out their experiments. To ensure that there was no interference between groups, the second group started their experiment immediately after the first group finished their lesson. The activities that the participants performed were as follows:

Identification of the Task: In this stage, participants received experiment brochures showing a problem to answer or a task to complete, and the tools available for use when performing their research. In this step, the researcher revealed the prior knowledge of participants and attempted to have them build a connection between their prior knowledge and the current research subject. These brochures were prepared by the researcher based on sections from the ADI brochures prepared by Hall and Sampson (2009) and Samson et al. (2011), after which they were submitted to expert opinion.

Generation of Data: Groups in this stage developed various methods for answering the research questions. The researcher visited each group individually, gave clues to help them start their experiment, and later asked questions about the methods they had developed.

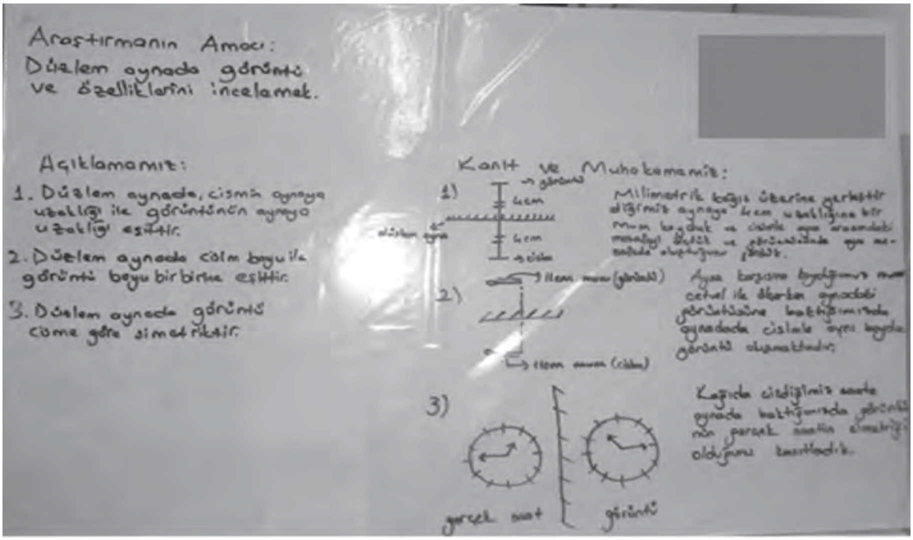
Participants in this stage had the opportunity to learn what kind of materials and data collection techniques to use according to the subject and the qualification of the research, as well as why some methods yield better results than others. They also experienced how to overcome problems they encountered in an experimental study (Samson & Gleim, 2009; Sampson et al., 2011).

Production of a Tentative Argument: After participants completed their research, they prepared a poster to share, defend, and present their ideas, as seen in Figure 1. This poster consisted of the names of the group members, purpose of the research, an explanation, and evidence and reasoning. The "Explanation" section of the argument gave an answer to the research question. The "Evidence" part of the argument included observations and measurements to present the validity of the explanation. Evidence consisted of numerical data (e.g. mass, time, temperature) or observation notes (e.g. the color has changed, it revealed a gas). The "Reasoning" part of the argument included justification of how the evidence supports the claim and whether it is a justifiable claim or not. In this stage, which was performed to draw attention to the importance of the argument, the participants understood that scientists should support an explanation with proper evidence and reasoning.

Interactive Argumentation Session: In this stage, participants debated over their research using the round-robin format in order to provide them with a critical perspective about the products (the arguments), the process (method) and the content (theoretical background) of their research. One member from each group remained on the table as a spokesperson to share the research they had carried out, the conclusion they had reached, the data they had collected and the ideas of their group. Other group members visited different tables to listen to and criticize the studies and arguments of other groups in the class. Later, participants went back to their own groups and conveyed the examples from the other groups to the group member who had remained. They gave demonstrations of those who used different methods, if any. In this way, participants had the opportunity to socially construct their knowledge.

According to Hall and Sampson (2009), this kind of experience helps students understand theoretical reviews, trials and expectations, and the beliefs that affect the problems that a scientist researches, as well as how to carry out research and interpret observations.

Creation of a Written Investigation Report: Participants prepared individual reports in line with the data they obtained and results they found.



[redacted]

Araştırmanın Amacı:
Düzlem aynada görüntü ve özelliklerini incelemek.

Açıklamamız:
1. Düzlem aynada, cismin aynaya uzaklığı ile görüntünün aynaya uzaklığı eşittir.
2. Düzlem aynada cisim boyu ile görüntü boyu birbirine eşittir.
3. Düzlem aynada görüntü cisme göre simetriktir.

Kanıt ve Muhakememiz:

Milimetrik kağıt üzerine yerleştirdiğimiz aynaya 4cm uzaklığına bir mum koyduk ve cisimle ayna arasındaki mesafeyi ölçtük ve [illegible] ayna [illegible] oluştuğunu gördük.

Ayna karşısına koyduğumuz mumu cetvel ile ölçerken aynadaki görüntüsüne baktığımızda aynada da cisimle aynı boyda görüntü oluşmaktadır;

Kağıda çizdiğimiz saate aynada baktığımızda görüntünün gerçek saatin simetriği olduğunu tespit ettik.

Figure 1: An example of a poster prepared by participants.

If practice time was not enough, reports were prepared outside of the class.

The Double-blind Peer Review: After participants prepared weekly reports, researchers and participants set up a time to evaluate the reports together. The reports were coded by the researcher and distributed back to the participants for evaluation by the group. The participants evaluated the reports according to the criteria developed by Sampson and Gleim (2009). Reporters and evaluators were kept confidential.

Due to time limitations, the two final stages of the ADI method, the revision process and reflective round-table discussion were not performed. Alternatively, after each experiment, the study results were evaluated by the class. The researcher examined the reports and gave feedback to the participants.

Control Group: Traditional Laboratory Method

As with the experimental group, prior to the laboratory course the researcher divided control group participants into heterogeneous groups of four people (two groups of three people) in terms of gender and success. Using this method, participants carried out their experiments in accordance with the laboratory handouts given every class hour. In their handouts, the name of the experiment, the research questions and objectives, tools to be used in the experiment, how to use them, and the activities of the experiment were given step by step. Participants wrote their reports outside the classroom after each experiment in accordance with the data obtained. After each experiment, the results were evaluated by the class. After the reports were delivered by the participants, the researcher examined the reports and gave feedback to the participants.

Data Collection

Data was collected by the researcher during the fall semester of the 2010-2011 academic year. Before starting the treatment, the Optical Achievement Test, Argumentative Scales, and SPST were administered both to the control group and the experimental group as a pretest, and afterwards as a post-test. Before starting the laboratory courses, the first researcher gave information to the participants about how the courses would be conducted during the lessons using the ADI method for the experimental group, and the traditional laboratory method for the control group. A sample report from the literature was given to the group using the ADI method to inform them about how to prepare their reports. Since the control group prepared their reports using the traditional method the previous year, they were not reminded about how to prepare their reports again. In the fall semester of the 2010-2011 academic year, the control group participated in the traditional laboratory course, while the experimental group participated in the ADI laboratory course. At the end of the treatment, post-tests were administered to both groups.

Data Analysis

To determine whether there was any significant difference between the control group and the experimental group regarding achievement levels and science process skills, the researchers implemented ANCOVA using the OAT and SPST scores of the participants. The primary purpose of the ANCOVA analysis was to examine if their assumptions had been met.

To determine whether there was any significant difference between the control group and the experimental group on their argumentativeness and argumentation levels, the researchers implemented the independent t-test using the AS and argumentation level (AL) scores of the participants.

To determine whether there was any significant difference between the control group and the experimental group regarding argumentation levels, and to study the argumentation level changes, reports that were written for the first, the middle (fourth), and last experiments of the treatment were examined in two stages. Firstly, the arguments written in the participants' reports were divided into components according to the Toulmin Model. There are six items in the Toulmin Model. The first three of them (data, claim, and warrants) form the basis of an argument (Kneupper, 1978; Toulmin, 1990). The other three items (rebuttals, backing and qualifiers) are subsidiary elements of the argument (Toulmin, 1990).

Descriptions of the items in the Toulmin Argument Model are as follows (Driver et al., 2000):

Data (D): The facts used to prove the claim.

Claim (C): The results presented based on the data.

Warrants (W): Logical statements that prove the relation between the claim and the data.

Backing (B): The main assertions that prove the warrants.

Qualifiers (Q): The statements that limit the argument and propose the conditions under which the argument is true.

Rebuttals (R): Counter-arguments or statements that specify the conditions when the argument might not be true.

In the second stage, written arguments (student reports) which had been resolved into their constituents in accordance with the description given in the Toulmin Model were put into stages based on the argumentation-level model developed by Erduran, Simon, and Osborne (2004). Assuming that the quality of the arguments including rebuttals was high, Erduran et al. (2004) created a framework consisting of five levels to assess the quality of argumentations. While creating these levels, they considered the argument model developed by Toulmin. These argument levels are given in Table 1.

Table 1
Analytical Framework Used for Assessing the Quality of Argumentation (Erduran et al., 2004; p. 928)

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.

In this study, the first four levels were used to determine the argumentation levels of student-written reports, because no student reports were in the fifth level.

Each level was scored by the researcher, and these scores were used while implementing the t-test. Levels were scored as 1 for Level 1, 2 for Level 2, 3 for Level 3, and 4 for Level 4. Levels were explained with detailed examples in the findings.

All individual student reports that were examined for understanding argumentation level were coded by the researcher. To ensure reliability, 48 data sets chosen randomly among all the data (24 from the experimental group, and 24 from the control group) were coded by a second coder.

The second coder was an expert working in the field of argumentation. After the two coders independently completed their encoding, student reports that were coded with different argumentation levels were examined and the reasons for these differences in coding were discussed by the coder. The researcher corrected items that were considered to be wrongly encoded according to the coding scheme based on the Toulmin Model. The inter-rater reliability of the coders was calculated using Cohen Kappa, and the Cohen Kappa coefficient was found to be .82.

Findings

Findings for the Optical Achievement Test

Prior to treatment, the average scores of the experimental and control groups showed a significant difference ($X_{experimental} = 10.76$, $X_{control} = 8.82$, $t_{77} = 2.39$, $p < .05$). Therefore, ANCOVA was used for the comparison of the OAT post-test scores of the two groups. After ANCOVA assumptions were met, descriptive statistics for the OAT post-test scores were investigated. Descriptive statistics for the OAT post-test scores are given in Table 2.

Table 2
Descriptive Statistics of the Optical Achievement Post-Test Mean Scores According to Group

Group	N	Mean	SD	Corrected Mean	Std. Error
Experimental	41	16.56	3.32	16.13	0.48
Control	38	12.08	3.51	12.54	0.50

The results of ANCOVA analysis for understanding the significance of the difference between the corrected OAT mean scores of the groups are presented in Table 3.

Table 3
ANCOVA Results of the Optical Achievement Post-test Mean Scores Corrected According to Optical Achievement Pre-Test

Source	Sum of Squares	df	Mean Square	F	*p*	eta²
OAT pretest scores	212.40	1	212.40	23.65	.00*	.237
Group	236.41	1	236.41	26.33	.00*	.256
Error	682.46	76	8.98			
Corrected Total	1291.04	78				

According to the ANCOVA results, when the OAT pretest scores of the experimental and control groups were analyzed, a significant difference between their corrected post-test scores ($F_{(1-76)} = 26.33$, $p = .00$) was found. The effect size of the applied method is $\eta^2 = .26$, which means 26% of the variance in the average OAT score of the two groups is described by this method. An effect size over .14 is statistically considered a large effect

(Peirce, Block, & Aguinis, 2004). Accordingly, it could be said that the ADI method has a big effect on the academic achievement of pre-service science teachers regarding geometrical optics.

Findings about Argumentative Scale

Prior to the treatment, in order understand whether there was a significant difference between the AS pre-test scores of the experimental and control groups, independent samples t-test was carried out. According to the t-test results, there was no significant difference between the pre-test scores of the groups ($X_{experimental} = 71.41$, $X_{control} = 67.76$; $t_{77} = 1.31$, $p > .05$). Therefore, the independent samples *t*-test was used to compare the AS post-test scores of the groups. The results are given in Table 4.

Table 4
Experimental and Control Groups' Argumentative Scale Post-Test T-test Results According to Group Mean Scores

Groups	N	X	SD	df	*t*	*p*
Experimental	41	75.35	10.41	77	1.56	.12
Control	38	71.81	9.83			

According to Table 4, it can be seen that there is no significant difference between the post-test scores of the experimental and control groups ($t_{(77)} = 1.56$, $p > .05$).

Findings about Science Process Skill Test

Prior to treatment, it was found that there was a significant difference between the pretest scores of the experimental and control groups ($X_{experimental} = 22.83$, $X_{control} = 17.32$; $t_{77} = 6.11$, $p < .05$). Therefore, ANCOVA was used to compare the SPST post-test results of the groups, and the SPST pretest scores were taken as covariates.

After ANCOVA assumptions were met, descriptive statistics for the SPST post-test scores were analyzed. Descriptive statistics for the SPST post-test scores are given in Table 5.

Table 5
Descriptive Statistics of Science Process Skills Post-Test Mean Scores According to Groups

Group	N	Mean	SD	Corrected Mean	Std. Error
Experimental	41	24.46	3.85	23.17	.57
Control	38	19.68	3.75	21.08	.60

To reveal any possible significant difference between the two groups, the ANCOVA results of the corrected SPST mean scores of the two groups are given in Table 6.

Table 6
ANCOVA Results of Science Process Skills Post-Test Mean Scores Winsorized According to Science Process Skills Pretest

Source	Sum of Squares	df	Mean Square	F	*p*	eta^2
SPST pretest scores	291.76	1	291.76	26.96	.00**	.26
Groups	58.44	1	58.44	5.4	.02*	.07
Error	822.64	76	10.82			
Corrected Total	40375	78				

When Table 6 is examined, it is seen that after the SPST pretest scores were controlled, there was a significant difference on the SPST post-test scores in favor of the experimental group ($F_{(1\text{-}76)} = 5.4$, $p = .023$). An effect size of .06 or higher is interpreted as a moderate effect (Perce et al., 2004). When the effect size is analyzed, a moderate effect for the method is seen ($\eta^2 = .07$), and ADI explains 7% of the variance in the scientific process skills of the participants.

Findings about Argumentation Levels

To examine the changes in the experimental group participants' argumentation levels, the reports written by the participants in the first experiment, the fourth experiment (middle) and the last experiment of the treatment were analyzed according to argumentation levels as developed by Erduran et al. (2004). The arguments produced by the participants were classified into four levels after analysis. These were Levels 1, 2, 3, and 4 respectively. Each level was described with respective examples. As a result of the examined test reports, development of the argumentation levels and mean scores of the argumentation levels of the experimental and control groups were then compared.

Level 1: This level includes only reports that contain a claim. In these reports there is no data, warrant, backing, or rebuttal units. The researcher scored this level as 1 point. An example of a Level 1 student report is given below.

Student 1 (S1): *combining our prior knowledge, we carried out research as a group. We learned that when someone gets close to a plain mirror the sight widens, and it narrows if someone walks away from the mirror* (the claim). *We realized that to see all of an object, the mirror should be half the length of the body* (the claim).

S1 revealed only the claims in the report he/she wrote after the experiment. He/she didn't use any data, warrant, backing or rebuttal related to the claim.

Figure 2: The illustration showing the evidence for S4's claims.

Level 2: This level includes reports that contain data, warrant and backing together with the claim, but without rebuttal. This level is scored with 2 points. An example of a Level 2 student report is given below.

S4: *We sought to prove that the image is flat and at the same time it is smaller than the object. We explained this in the following way:*

One of our friends set the exact point of view so that the chin and the forehead touched the top and the bottom of the convex mirror. In this case, the full length of the face image would be equal to the length of the mirror. Our friend was 35 cm away from the mirror (the data). *And although his/her real face is 16 cm in length, the image in the convex mirror (4cm) was smaller* (the data). *As our friend approached the mirror, we saw that the length of the image grew.*

In this case, Betul sees her image flat and it grows as she gets closer to the convex mirror (the claim).

In this example, the student sought to prove the claim through numeric data and the illustration shown in Figure 2. Rebuttal is not used in the report.

Level 3: This level includes reports that contain data, warrant and backing, as well as weak rebuttals. Weak rebuttals are rebuttals that are made without using any evidence. The researcher scored this level with 3 points. An example of a Level 3 student report is given below.

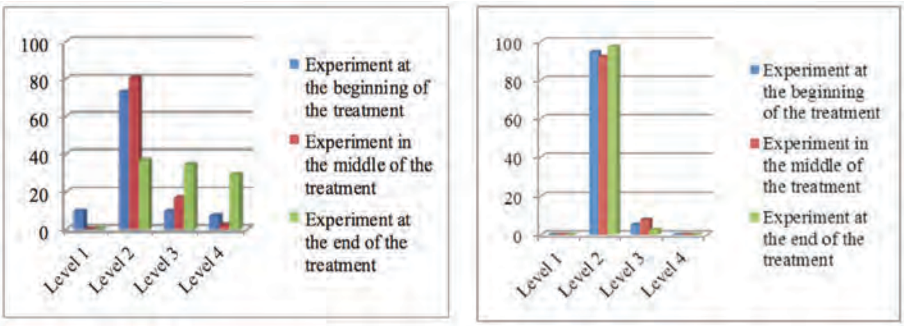
S8: *The prediction that Abrek's student made is correct* (the claim). *But we can also get the color of other teams* (weak rebuttal). *We got Fenerbahce and Galatasaray by using their colors.*

Here, the student used the rebuttal that he/she would be able to get the colors of other teams, but didn't put forward any evidence about how he/she can get those colors.

Level 4: This level includes reports that use clearly defined rebuttals. A clearly defined rebuttal is a rebuttal that uses evidence that includes data, warrant, and backing (Erduran, 2008). The researcher scored this level with 4 points. An example of a Level 4 student report is given below.

S9: *As a result, Elbruz is right* (the data). *Trabzonspor has red and blue colors* (the data). *But we can also*

Figure 3: The illustration showing the evidence for S9's claims.



a) Experimental group

b) Control group

Figure 4: Percentage distribution of experimental group (*a*) and control (*b*) group participants' argumentation levels at the beginning, middle and end of the treatment.

see the colors of other teams (rebuttal). *Brazil has yellow-green, Galatasary has yellow-red and Gana has yellow-red-green* (the data).

When this example is compared to S8's example of a Level 3, it can be seen that as in S8's example, the colors of the other teams besides Trabzonspor could also be surmised, but this time the rebuttal is supported through data and the illustration given in Figure 3.

Changes in Argumentation Level

Argumentation levels of the experimental group and the control group at the beginning, in the middle and at the end of the treatment are given in Figure 4.

According to the figure, it is seen that most of the participants in the experimental group are at level 2 in the beginning and middle of the treatment, and the levels of the participants appear to be evenly distributed among Levels 2, 3 and 4 at the end of the treatment. On the other hand, in the control group, almost all participants are at Level 2 in the beginning, middle and end of the treatment. There are no participants at Level 4.

To show whether there were any significant differences between the argumentation-level mean scores of the experimental group and the control group in the beginning, middle and end of the treatment, the independent samples t-test was conducted. The results are given in Table 7.

Table 7
T-test Results by Group for Argumentation Level Mean Scores of the Experimental Group and the Control Group at the Beginning, Middle and End of the Treatment

Experiment	Group	N	X	SD	df	*t*
The first	experimental	41	2.15	0.69	77	0.8
	control	38	2.05	0.23		
The middle	experimental	41	2.22	0.47	77	1.6
	control	38	2.08	0.27		
The last	experimental	41	2.93	0.82	77	6.61*
	control	38	2.03	0.16		

*p < .01

According to Table 7, it was found that in the beginning ($t_{(77)}$ = 0.8, p = .43), and the middle ($t_{(77)}$ = 1.6, p = 0.12) of the treatment there was no significant difference between argumentation level scores of the experimental and control group participants. However, at the end of the treatment, the argumentation-level scores of the experimental group were found to be significantly higher than the scores of the control group participants ($t_{(77)}$ = 6.61, p = .00)

Discussion

Discussion on Optical Achievement Test Results

According to the findings of the study, in the General Physic Laboratory III class for sophomores in which the experimental participants used the Argument-driven Inquiry Method, a statistically significant difference was identified between the academic achievement of the experimental group and the control group in favor of the experimental group. Based on these results, it can be said that the laboratory instruction method of Argument-Driven Inquiry increased the academic achievement of the pre-service science teachers. Similar results have been listed in the literature. For example, Science Writing Heuristic (SWH), similar to ADI, provided university students with higher scores in chemistry tests (Akkuş et al., 2007; Burke et al., 2005).

The ADI method, especially the argumentation phase, is thought to have a major effect on increasing success in this study. Participants discussed the arguments they produced. During the discussion they had knowledge about the ideas, and they had a chance to fix any errors they made. When participants figured out that other groups in the class were applying different methods, they wanted their friends to show them what they were using, and they had a chance to observe alternative methods. Cross et al. (2008) appear to support this finding. In this study, it was found that argumentation and collaborative group work influence student learning and success in science. The conceptual knowledge test scores of the experimental group who took argumentation education increased in Zohar and Nemet (2002).

Discussion on Argumentative Scale (AS) Results

The results of the Argumentative Scale (AS) show that there was no significant difference between the post-test scores of the experimental group and the control group. According to these results, it can be claimed that laboratory training based on the Argument-driven Inquiry method does not change pre-service science teachers' AS scores when compared to the traditional laboratory method. In the fourth step of the ADI method, the argumentation phase, participants in the experimental group conducted discussions in round-robin fashion to listen to and criticize the study and arguments of the participants in other groups. However, participants in the traditional laboratory method instruction implemented experiments according to the handout they had and they didn't participate in any discussion activity. When one considers this situation, since the experimental group participants participated in discussion activities after each experiment, it might be expected that there would be a difference between their AS scores and the AS scores of the control group. However, the lack of any significant difference between the groups with their AS scores can be due to a single academic year's worth of laboratory activities not being enough to change argumentativeness (Osborne, Erduran, & Simon, 2004).

In contrast to the results of this study, it has been claimed in the literature that argumentation increases student argumentativeness (Kaya & Kılıç, 2008; Rancer et al., 1997). Rancer et al. (1997) implemented a program for about seven days including sections in which students participated in an argument on two subjects. They found that after the implementation, student argumentativeness had increased. Kaya and Kılıç (2008) found that after a semester of argumentation activities, student argumentativeness had increased when compared to their levels before the study.

One of the reasons for not reaching similar results with other studies in the literature can be explained with the sampling differences. Kaya and Kılıç (2008) worked with seventh and eighth grade students, while Rancer et al. (1997) studied seventh grade students between the ages of 11 and 13. In their study on the relation with argumentativeness to age and level of education, Scullery and Scullery (2003) found that the argumentativeness decreased with increasing age. Lack of a significant difference between the groups at the university level suggests that age might be a factor in this case. Besides, Kaya and Kılıç carried out their research with 23 seventh grade and 24 eight grade students (approximately half of the samples in this research). A difference in method application might be the reason for different results. Kaya and Kılıç used a single group pretest and post-test design. Student argumentativeness was examined from the beginning to the end of the treatment. Also in those studies, student participation time in argumentation was longer than the time in this study. While participants engaged in the argumentation phase for only 20 minutes in this study due to total class time limitations, for Kaya and Kılıç, students participated in the argumentation phase for up to 47 minutes using peer, small group, and classroom discussions. In Rancer et al., the argumentation process included the entire duration of the class. In both studies, the use of socio-scientific issues may have influenced argumentativeness. The development of argumentation with scientific subjects is more difficult than with development in socio-scientific issues (Osborne et al., 2004), because in socio-scientific issues, students can use knowledge gained from daily life experiences. Therefore, they can express themselves more comfortably with a tendency to argue about controversial issues. With scientific issues, however, more specific information is needed and the tendency to avoid arguments can be higher.

It was found that while there was a significant difference between the experimental and control groups' academic success, scientific process skills and argumentation levels, there was no such difference with their argumentativeness. A relationship between argumentativeness and the quality of an argument could be expected, but in this study the quality of the arguments was determined solely by the quality of student-written arguments. Participants who had refrained from verbal discussion during the argumentation phase of the ADI had the chance to consider the arguments made by others in their group, and the opportunity to use this in their written arguments could be a factor in this case. While participants avoided verbal arguments, they tended to express their arguments in written forms.

Discussion on the Science Process Skills Test Results

According to the Science Process Skills Test (SPST) results, there was a significant difference between the experimental and control groups on their SPST post-test scores in favor of the experimental group. Considering these findings, it was found that laboratory instruction based on the Argument-driven Inquiry Method was more effective than the traditional laboratory method for increasing the science process skills of pre-service science teachers. In their study investigating the effects of laboratory techniques on pre-service science teachers' science process skills, Myers and Dyer (2005) found that students taking inquiry laboratory instruction got higher science process skill scores than students taking traditional laboratory instruction. In Tümay and Köseoğlu (2011), it was found that after an argument-based chemistry class, pre-service chemistry teachers thought argumentation improved various skills including inquiry and scientific thinking. Duru et al. (2011) found that inquiry based laboratory practices increased the ability of students to use science processes. Günel et al. (2012) stated that in the argument-based science learning approach, students construct knowledge using a research-inquiry based learning environment in which they ask questions, form claims and prove them through clues. They claim that this approach constructs a more effective learning environment by increasing student participation in the learning process. According to Marbie and Baker (1996), there was a link between the experimental teaching strategies and the development of science process skills. Welch (1981) stated that while gains, principles and learning facts are at the forefront of the traditional laboratory method, skills like problem solving, scientific thinking and inquiry are not addressed in this method. Aktamış (2007) states these kinds of experiments do not help students gain science process skills. In the ADI method, students design their own experiment, hypothesize on their own, and create graphs and shapes in line with the results they find. On the other hand, since students who used the traditional laboratory method carried out their experiments according to the handouts given to them, the theoretical information from the handouts directed them to the result. They are even given the variables for the graph they are supposed to prepare, and students are not provided with a chance to explore their own ideas. In the ADI method, students again observe that data is presented in various figures and graphs, and during the discussions they seek the differences in interpretation of data. During the "identification of the task" phase in the ADI method, students are given a brochure which includes a short summary, research question, and a problem to solve or a task to complete. In this phase, students are provided with case studies in some brochures, and students construct their research problems according to these cases then carry out their experiments. In the argumentation phase, students asked each other questions, constructed claims, shared their acquired data, and interpreted these together. They constructed knowledge through inquiry-based practices. Burke et al. (2005) stated that with the practices of the argumentation phases, students find opportunities to construct knowledge through reasoning and argumentation, as well as experience their science processes. It is also thought that the phases of the ADI method contribute to the development of the science process skills of students.

Discussion on the Argumentation Level

According to the findings regarding changes in argumentation levels, most participants of the experimental and control groups were at Level 2, but it was found that throughout the study the number of the participants in Level 3 and Level 4 increased. While 9.8% of the participants in the experimental group were in Level 1 at the beginning of the treatment, there were no Level 1 students in the middle or at the end of the treatment. Additionally, in the middle of treatment, the number of the Level 2 participants increased when compared to the beginning of the treatment. It can be stated that one reason for this result might be related to the fact that participants who were using only claims in their reports in Level 1 started using constituents that would support and prove their claims in the middle of the treatment, thereby proceeding to Level 2. Rebuttals are the indicators of high-quality argumentation, and discussions using rebuttals have a higher level of argumentation (Aufschnaiter et al., 2008; Erduran et al., 2004). An increase in the number of Level 3 and 4 participants using rebuttals indicates the increase of argumentation levels of the experimental group participants. It can also be said that as the study continued, the participants adopted different point of views because participants started using rebuttals and addressing situations with invalid claims.

On the other hand, in the control group almost all participants were in Level 2 at the beginning, middle and end of the treatment. This shows that participants in this group have low argumentation levels. As almost all participants stayed in Level 2 throughout the study, this might indicate that they didn't adopt different points of view and kept their ideas in the same form. As Aufschnaiter et al. (2008) stated, the absence of rebuttals in arguments indicates no change in ideas by keeping them constantly the same.

A similar result is noted in Osborne et al. (2004). In this study, the majority of student arguments were Level 2 both at the beginning and at the end of the study. There was a decline in the number of Level 2 participants when the end of the study was compared to the beginning, and there was an increase in the number of Level 3 and higher participants when compared to the beginning of the study. Considering the fact that Osborne et al. worked with high school students in their study, it can be claimed that argumentation levels for different age groups are the same. Samson et al. (2010) also found that students produced high quality written arguments after the treatment they carried out with the university students.

Traditional reports consist of the following sections: name of the experiment, objective of the experiment, procedure, data, statistics, results, and commentary. Control group participants presented an explanation (claim) answering the research questions during the results section and they presented the obtained data in the data section of their reports. For the experimental group participants, there was a section for answering the question “What is your argument?” in which participants presented a good explanation for precisely answering the research question. They asserted valid and reliable evidences to support their explanation. Participants in this section also mentioned their rebuttals, referring to situations where their findings would not be valid. As a result of this method, different report formats might be a determining factor for the quality of student argumentation. It can be assumed that in the argumentation phase, experimental group participants had a chance to criticize arguments that other groups had put forth. They found an opportunity to refute each other's ideas, so this eased the process for writing their reports and helped them make their rebuttals. The fact that there was no significant difference between the argumentation levels until the end of the treatment might prove that argumentation levels do not change in a short period. It took an entire semester to increase, and a long time is required to effect a change in argumentation levels (Osborne et al., 2004).

The studies of Zohar and Nemet (2002) indicate similar results. The current study also shows that after argumentation training, students' argumentation skills had statistically changed. In Walker et al. (2010), it was found that after the study applied the ADI method, there was a significant difference between students using the ADI method and those using the traditional laboratory method regarding their use of evidence and reasoning skills. But unlike these studies, Osborne et al. (2004) found no significant difference between the groups at the end of the study in terms of argumentation levels.

According to these results, it can be claimed that laboratory training using the argument-driven inquiry method is more effective than the traditional laboratory method in increasing argumentation levels, in other words the argumentation quality of science teacher candidates. Laboratory training that uses the argument-driven inquiry method also needs to be performed longer term in order to observe those changes.

Suggestions

It is necessary to use new teaching models different from the traditional methods to simultaneously include science content, scientific processes, and social norms in laboratory courses (Sampson at el., 2011). It was found that the ADI method is more effective than the traditional laboratory method on increasing academic achievement, scientific process skills and argumentation levels. In this case, the ADI method might be used in laboratory courses instead of the traditional laboratory method. Additionally, teachers might include the ADI method not only in the lab but also in other classes.

A longer length of time is required to use student-centered methods in laboratory courses when compared to teacher-centered methods. The time allocated for laboratory courses in the current program was not enough for these kinds of methods. In this study, peer review and report writing phases were performed outside of the regular teaching hours. Therefore, extension of the time allocated for laboratory courses is recommended for the prepared programs. Additionally, there was no difference with the argumentativeness of the participants when compared to the traditional laboratory method. It is thought this is because the time allocated in a semester for student participation in the argumentation process and laboratory activities is not enough to change the tendency towards discussion. Increasing the time allocated for the argumentation phase in the ADI method and carrying out studies over a longer length of time in a single academic semester might be suggested. Additionally, the difference found in the literature between students of various age categories in regard to their tendency to argue controversial issues suggests that age might also be a factor in this case. The age factor can also be examined by working with students from various levels of education like different classes from the primary, secondary and university levels.

Developing argumentation in scientific issues is harder than developing argumentation in socio-scientific issues (Osborne et al., 2004). In this study, argumentativeness and the level of argumentation in scientific issues were examined. Using the ADI method, studies focusing on socio-scientific issues could be carried out.

In this study, the argumentation levels of participants were identified by examining written arguments from the individual reports they prepared after each experiment. Those reports were analyzed according to the model of argumentation level developed by Osborne et al. (2004). Video and audio recordings of the students' verbal arguments during the "argumentation phase" in the ADI application might be examined, and student arguments can be evaluated from either an epistemic perspective or in terms of the scientific knowledge content and its defense by using various analytical methods.

References

Anderson, R. D. (2007). Inquiry as an organizing theme for science curricula. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 807-830). Mahwah, NJ: Lawrence Erlbaum Associates.

Akkuş, R., Günel, M., & Hand, B. (2007). Comparing an inquiry-based approach known as the science writing heuristic to traditional science teaching practices: Are there differences? *International Journal of Science Education, 29*(14), 1745-1765.

Aktamış, H. (2007). *Fen eğitiminde bilimsel süreç becerilerinin bilimsel yaratıcılığa etkisi: ilköğretim 7. sınıf fizik ünitesi örneği* (Doctoral dissertation, Dokuz Eylül University, Izmir, Turkey). Retrieved from https://tez.yok.gov.tr/UlusalTezMerkezi

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.

Aydoğdu, B., & Ergin, Ö. (2008). Fen ve Teknoloji Dersinde kullanılan farklı deney tekniklerinin öğrencilerin bilimsel süreç becerilerine etkileri. *Ege Eğitim Dergisi, 9*(2), 15-36.

Burke, K. A., Hand, P., Poack, J., & Greenbowe, T. (2005). Using the science writing heuristic. *Journal of College Science Teaching*, *35*(1), 36-41.

Cobern, W. W., Schuster, D., Adams, B., Applegate, B., Skjold, B., Undreiu, A. … Gobert, J. D. (2010). Experimental comparison of inquiry and direct instruction in science. *Research in Science & Technological Education, 28*(1), 81-96.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London: Routledge Falmer.

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.

Demircioğlu, T., & Ucar, S. (2012). The effect of argument-driven inquiry on pre-service science teachers' attitudes and argumentation skills. *Procedia-Social and Behavioral Sciences, 46*, 5035-5039.

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

Duru, M. K., Demir, S., Önen, F., & Benzer, E. (2011). Sorgulamaya dayalı laboratuar uygulamalarının öğretmen adaylarının laboratuar algısına tutumuna ve bilimsel süreç becerilerine etkisi. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, *33*, 25-44.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*, 39-72.

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. (2008). Methodological foundations in the study of argumentation in science classrooms. In S. Erduran & M. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 47–69). Dordrecht, Netherlands: Springer.

Feyzioğlu, B., Demirdağ, B., Ateş, A., Çobanoğlu, İ., Altun, E., & Akyıldız, M. (2011). Laboratuvar uygulamalarına yönelik öğrenci görüşleri: İzmir ili örneği. *İlköğretim Online, 10*(3), 1208-1226.

Geban, Ö., Aşkar, P., & Özkan, İ. (1992). Effects of computer simulated experiments and problem solving approaches on high school students. *Journal of Educational Research, 86*, 5-10.

Günel, M., Kıngır, S., & Geban, Ö. (2012). Argümantasyon tabanlı bilim öğrenme (ATBÖ) yaklaşımının kullanıldığı sınıflarda argümantasyon ve soru yapılarının incelenmesi. *Eğitim ve Bilim, 37*(164), 316-330.

Hall, C. B., & Sampson, V. (2009). Inquiry, argumentation, and the phases of the moon: Helping students learn important concepts and practices. *The Science Scope, 32*(8), 16-21.

Hiatt, D. B. (1986, April). *Post-Sputnik educational reform era: To dream the impossible dream*. Paper presented at the Annual Meeting of the California Educational Research Association, Los Angeles.

Harlen, W. (1999). Purposes and procedures for assessing science process skills. *Assessment in Education, 6*(1), 129-144.

Hoffstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for twenty-first century. *International Journal of Science Education, 88*, 28-54.

Infante, D. A., & Rancer, A. S. (1982). A conceptualization and measure of argumentativeness. *Journal of Personality Assessment, 46*, 72-80.

Jiménez-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. 3-27). Dordrecht, Netherlands: Springer.

Jiménez-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.

Kaya, O. N., & Kılıç, Z. (2008). Development of elementary school students' argumentativeness in science courses. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi (KEFAD), 9*(1), 87-95.

Kelly, G. J., & Takao, A. (2002). Epistemic levels in argument: An analysis of university oceanography students' use of evidence in writing. *Science Education*, *86*(3), 314–342.

Kelly, G. J., Druker, S., & Chen, C. (1998). Students' reasoning about electricity: Combining performance assessments with argumentation analysis. *International Journal of Science Education, 20*(7), 849–871.

Kim, H., & Song, J. (2005). The features of peer argumentation in middle school students' scientific ınquiry. *Research in Science Education, 36*(3), 211-213.

Kneupper, C. W. (1978). Teaching argument: An introduction to the Toulmin model. *College Composition and Communication*, *29*(3), 237-241.

Lunetta, V. N., Hoffstein, A., & Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory and practice. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 75-102). Mahwah, NJ: Lawrence Erlbaum Associates.

Marbie, R., & Baker, M. (1996). A comparison of experiential instructional strategies upon the science process skills of urban elementary students. *Journal of Agricultural Education, 37*(2), 1-7.

Myers, B. E., & Dyer, J. E. (2006). Effects of investigative laboratory instruction on student content knowledge and science process skill achievement across learning styles. *Journal of Agricultural Education, 47*(4), 52-63.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argumentation in the pedagogy of school science. *International Journal of Science Education*, *21*(5), 553–576.

Nussbaum, E. M., & Sinatra, G. M. (2003). Argument and conceptual engagement. *Contemporary Educational Psychology, 28,* 384-395.

Okey, J. R., Wise, K. C., & Burns, J. C. (1982). *Test of integrated process skills (TIPS II).* Athens: University of Georgia, Department of Science Education.

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.

Pierce, C. A., Block, C. A., & Aguinis, H. (2004). Cautionary note on reporting eta-squared values from multifactor ANOVA designs. *Educational and Psychological Measurement*, *64*(6), 916-924.

Rancer, A. S., Whitecap, V. G., Kosberg, R. L., & Avtgis, T. A. (1997). Testing the efficacy of a communication training program to increase argumentativeness and argumentative behavior in adolescents. *Communication Education, 46*(4), 273- 286.

Sampson, V. (2009, April). *The impact of Argument-Driven Inquiry on three scientific practices.* Paper presented at Annual International Conference of the National Association of Research in Science Teaching (NARST), Garden Grove, CA.

Sampson, V., & Gleim, L. (2009). Argument-driven inquiry 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 inquiry 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.

Sampson, V., Walker, J., Dial, K., & Swanson, J. (2010, March). *Learning to write in undergraduate chemistry: The impact of argument-driven inquiry.* Paper presented at the 2010 Annual International Conference of the National Association of Research in Science Teaching (NARST), Philadelphia, PA.

Schen, M. S. (2007). *Scientific reasoning skills development in the introductory biology courses for undergraduates* (Doctoral dissertation). Retrieved from https://etd.ohiolink.edu/!etd.send_file?accession=osu1187063957&disposition=inline

Schullery, N. M., & Schullery, S. E. (2003). Relationship of argumentativeness to age and higher education. *Western Journal of Communication, 67*(2), 207-223.

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International Journal of Science Education, 28*(2/3), 235–260.

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

Tümay, H., & Köseoğlu, F. (2011). Kimya öğretmen adaylarının argümantasyon odaklı öğretim konusunda anlayışlarının geliştirilmesi. *Türk Fen Eğitimi Dergisi*, *8*(3), 105-119.

Walker, J. P., Sampson, V., Grooms, J., Anderson, B., & Zimmerman, C. (2010, March). *Argument driven ınquiry: An instructional model for use in undergraduate chemistry labs*. Paper presented at the 2010 Annual International Conference of the National Association of Research in Science Teaching (NARST), Philadelphia, PA.

Watson, J. R., Swain, J. R. L., & McRobbie, C. (2004). Students' discussions in practical scientific inquiries. *International Journal of Science Education, 26*(1), 25-45.

Welch, W. W., Klopfer, L. E., Aikenhead, G. S., & Robinson, J. T. (1981). The role of inquiry in science education: Analysis and recommendations. *Science Education, 65*, 33-50.

Wellington, J., & Osborne, J. (2001). *Language and literacy in science education.* Philadelphia, PA: Open University Press.

Yağbasan, R., & Kanlı, U. (2008). 7E modeli merkezli laboratuvar yaklaşımının öğrencilerin bilimsel süreç becerilerini geliştirmedeki yeterliliği. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, *28*(1), 91-125.

Zeidler, D. L. (1997). The central role of fallacious thinking in science education. *Science Education, 81*, 483-496.

Zohar, A. (2008). Science teacher education and professional development in argumentation. In S. Erduran & M. P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: perspectives from classroom-based research* (pp. 245-268). Dordrecht, Netherlands: Springer.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, *39*, 35–62.