ARTICLE

The Laboratory as a Vehicle for Enhancing Argumentation Among Pre-Service Science Teachers

Naim Najami[1] · Muhamad Hugerat[1] · Fattma Kabya[1] · Avi Hofstein[1]

Abstract

A study was conducted on how scientific laboratory lessons contribute to building arguments, both cognitively and socially. The population consisted of 12 second-year pre-service teachers at a college of education specializing in the teaching of biology and chemistry in middle school. The study examines the nature of the arguments students raise when conducting a laboratory experiment and how conducting an experiment contributes to curricular science teaching in the laboratory. For this purpose, a number of methodologies were used: observations, experiment reports, and statistical analysis. The findings showed that groups conducting an open-ended experiment made more claims in their discourse than did groups that conducted a confirmatory experiment and that the level of argumentation in the open-ended experiment group was higher than in the confirmatory experiment group. Interestingly, despite the significant difference in the level and quantity of arguments in the discourses in the two types of experiments, no significant difference was found regarding the level of argumentation in the two groups' experiment reports. However, students confirmed the great importance of scientific laboratory lessons to the development of argumentative thinking, meaningful learning, and the students' relationship with their classmates.

Keywords Argumentation · Meaningful learning · Higher-order thinking · Pre-service teachers · Science education

1 Introduction

The task of the teacher is teaching. Teachers must be prepared to learn about new initiatives by combining different elements with knowledge about the benefits and drawbacks of the teaching process (Jaspers et al. 2014). During the teaching process, teachers gain experience

---

✉ Muhamad Hugerat
  muha4@arabcol.ac.il

[1] The Academic Arab College for Education in Israel, 22 Hahashmal St., P.O. Box 8349, 33145 Haifa, Israel

in constructing and promoting new methods, using various means for transmitting knowledge to students in as interesting a way as possible (Hobson et al. 2009).

In a document that defines standards in education [National Research Council (NRC) 1996], the word "inquiry" has two meanings (Bybee 2000). One meaning of inquiry is to understand a topic by providing students with an opportunity to construct concepts and mental structures that enable them to understand the phenomena that they experience. The other meaning is the potential to develop many important high-order learning skills such as asking questions, developing critical thinking, problem-solving, and developing metacognitive and argumentation skills (Kipnis and Hofstein 2008; Hofstein and Kind 2012). Moreover, the laboratory provides support for high-order learning inquiry skills that include observing, planning an experiment, asking relevant questions, hypothesizing, and analyzing the experimental results (Hofstein et al. 2004).

Improving science instruction by promoting and supporting the use of inquiry-based instruction is for moving away from textbook and lecture-style teaching in an effort to better align science learning with the practices of science (National Research Council 2013). The NRC suggested that inquiry-based instructions include students who learn how to: collect, use, and interpret data; make claims using evidence, and discuss science as debate to support claims with evidence from data. Further, previous research supports the use of hands-on activities as a component of inquiry-based instruction (National Research Council 2013; Therrien et al. 2014). Even with the suggestions from the NRC and NGSS, teaching practices related to inquiry-based instruction can vary widely.

Few studies have been conducted mainly during pre-service teacher's development programs. McDonald (2014) found that all five participants engaged in quality argumentation in socio-scientific tasks, with the majority of participants producing high-quality arguments (McDonald 2014). The construction of an argument can be described as a kind of discourse through which claims of knowledge are constructed separately and interconnected, and evaluated in light of empirical or theoretical evidence (Erduran et al. 2015).

Toulmin (1958) claimed that the construction of arguments is a human behavior associated with social situations. Kuhn (1991) expanded the concept's social aspect and distinguished between two types of arguments: (1) rhetorical arguments, which are meant to convince someone else that something is true, whereas (2) dialogical arguments, which are created through discourse among participants with different opinions. Argument construction enables students to develop an understanding of the creation of the world of knowledge, especially the evolution of scientific knowledge. The construction of knowledge through group discourse is an example of socio-cultural constructivist learning as described by Vygotsky (1978).

To construct a well-founded and reasoned argument, many studies (Erduran et al. 2004; Erduran 2018; Katchevich et al. 2013; Katchevich et al. 2014) have used Toulmin's (1958) model, according to which an argument contains the following components: claim, data, and warrant, the latter constituting a connection between the former two. A basic claim should be based on data. The warrant explains the connection between a claim and data and should convince one to accept the claim. A higher-level claim contains a theoretical basis or backing, a qualifier, or a rebuttal. The evaluation of claims, based on Toulmin's model, involves the structural aspect: the claim's components and their interrelations. The stress is on basing the claim on data and explaining why the data support the claim. In the case of an argumentative discourse, the element of refutation is also expressed.

Osborne et al. (2004) proposed a number of strategies involving activities to promote argumentation skills. One strategy is to expose students to a number of claims on a topic in

science, which students are asked to accept or reject using appropriate arguments. In another strategy, students are exposed to two competing theories that explain a phenomenon and provide evidence associated with the theories. Students are then asked to construct arguments using a structured outline that includes one of the following: (1) students are asked leading questions, (2) they are asked to predict the results of a certain experiment and to justify their prediction; they then observe the experiment and are asked to explain the result, or (3) they are asked to plan an experiment, carry it out, and discuss its results (Katchevich et al. 2013). The present study is based on Toulmin (1958) for argument construction (see Fig. 1).

Inquiry-based teaching was found to be an appropriate strategy (Wilson et al. 2010). Here, we clarify science laboratory activities as learning experiences in which preservice teachers interact with materials to observe and better understand the natural world. Note that assessing the educational effectiveness of the laboratory and its related learning skills requires distinguishing between the different modes of instruction, namely, the nature of the experiments in which the students are involved. Upon examining the type of activities, it was found that formulating arguments is central and significant in developing and conducting science activities. Consequently, it is reasonable to assume that imparting the meaning of scientific content and the essence of developing a scientific concept would be a way to formulate arguments (Erduran et al. 2004; Hofstein et al. 2008). The skills involved in performing open-ended and confirmatory inquiry experiments are listed in Table 1 (see, for example, Hofstein and Kind 2012). Inquiry through lessons in the laboratory in which open-ended inquiry takes place serves as an excellent learning environment for this kind of activity (Ozdem et al. 2013). The laboratory activities may consist of a laboratory testing the correlation between two variables, fieldwork to collect as many findings as possible, or laboratory work to identify materials (Gott and Duggan 2007).

Sampson and Gleim (2009) proposed the "argument-driven inquiry" model for laboratory learning, which aims at enabling biology teachers to use inquiry laboratories as part of the curriculum, where the emphasis is on understanding concepts in biology, critical thinking, and argument construction, as the way to build knowledge and confirm its validity. Tien and Stacy (1996) found that the guided inquiry group was better able to assess the evidence from the studies to which they had been exposed, and that their explanations of the findings were better founded.
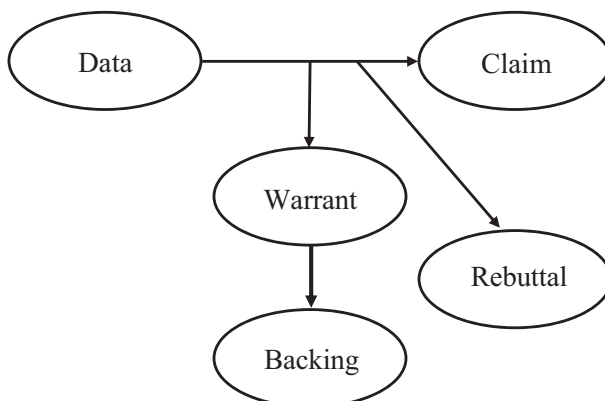

Fig. 1 Toulmin's argument pattern (Toulmin 1958)

Table 1 Skills involved in performing different types of laboratory activity

Skill	Open-ended	Confirmatory
Performing the experiment according to the teacher's instructions	+	+
Posing questions	+	
Constructing research questions	+	
Formulating hypotheses	+	
Caring out the experiment proposed by the students	+	
Arranging and organizing the results	+	+
Analyzing the results	+	+
Drawing conclusions	+	+
Summarizing the course of the experiment	+	+

2 Methodology

2.1 Objectives and Research Questions

The main objective of the present study was to expose classroom dialog and the process of argument construction to students of science education, in the context of the laboratory, during the discussion that develops while conducting experiments and afterwards. This aim resulted in the following research questions:

1. What are the discourse characteristics of preservice students during an open-ended inquiry experiment and a confirmatory experiment?

2. What are the argumentation components in the discourse among students of education during an open-ended inquiry experiment and a confirmatory experiment?

2.2 Research Design

The research design refers to the pilot study, population and the research procedures. The research method is mainly based on the use of qualitative tools. Some of the qualitative findings were analyzed quantitatively. The qualitative approach enabled us to describe in detail the phenomena and processes that occurred in the laboratory and that are related to constructing arguments. Quantitative analysis of the qualitative findings enabled us to describe the magnitude of the phenomena that we identified, with the goal in mind of comparing the different types of experiments, namely, the open-ended experiment vs. the confirmatory one.

2.3 Research Population

The research population consisted of 12 sophomore B.Ed. preservice teachers learning science education and specializing in biology and chemistry for middle school. The aim of the science education is to help the science preservice teachers to improve and develop their knowledge and skills in different aspects of science, to provide them with the relevant pedagogy in their work, and to equip them with the necessary skills and knowledge that are required in their career as science teachers and leaders in the education system in general, and in their field at schools in particular.

2.4 Research Tools

The research tools consisted of the following:

(1) Observing laboratory lessons and recording the discourses in them, (2) laboratory reports, and (3) interviews with students at the end of every experiment.

The observations focused on the spontaneous learning-related discourse that developed during the experiments. To each group of four students (totaling three groups) who conducted an experiment together, a recording device was attached to document the argument components. Then the levels of argument components used during the two types of experiments, according to Toulmin's (1958) model, were compared. The prior instructions required students to formulate a reasonable hypothesis, to analyze the results, and to write up the conclusions, in accordance with Chi (1997). In addition to the recording, the researcher also documented in writing the following aspects in the development of classroom discourse: the seating arrangement, cooperation among members of the group, and student participation in the developing discourse. All students conducted both types of experiments: confirmatory and open-ended inquiry experiments. The study was conducted in two stages.

2.4.1 Stage One: a Confirmatory Experiment on Respiration in Yeast

Students were given an instruction sheet:

Prepare dough in two bowls:

1) Form a ball of dough from 2½ teaspoons of flour, six teaspoons of water, and one teaspoon of yeast.
2) Form a ball of dough from 2½ teaspoons of flour and six teaspoons of water.
3) Place an equal amount of warm water (30 °C) in two glasses.
4) Place the dough ball from bowl no. 1 into glass no. 1—dough with yeast.
5) Place the dough ball from bowl no. 2 into glass no. 2—dough without yeast.

At first, both balls sink. However, after a short time, the ball of dough with yeast rises to the top, whereas the ball of dough without yeast remains at the bottom.

2.4.2 Stage Two: an Open-Ended Experiment on Respiration in Yeast

Students are given only equipment and materials, with no instruction sheet:

1) Form a ball of dough from flour, water, and a teaspoon of yeast.
2) Form a ball of dough from flour, water, a teaspoon of yeast, and sugar.
3) Place an equal amount of warm water (30 °C) in two glasses.
4) Place the dough ball from bowl no. 1 into glass no. 1 (− sugar).
5) Place the dough ball from bowl no. 2 into glass no. 2 (+ sugar).

At first, both balls sink. However, after a short time, the ball of dough with sugar rises to the top, whereas the ball of dough without sugar remains at the bottom.

We examined the processes of argument construction during the experiment and documented the presence or absence of the following basic argument components: (1) claim, (2) data, (3) warrant, (4) backing, and (5) rebuttal, in accordance with Toulmin's model (Fig. 1).

The research design presented above clearly shows the relationship between the research design and the research questions mentioned. Hence, the research design answers the research questions and this is the aim of the study.

2.5 Analysis of Laboratory Reports

The laboratory reports were written by a representative of each group and were collected at the end of the laboratory lesson. The reports' sections, the hypotheses they contained, and the conclusions were analyzed according to the following criteria: (1) Identifying the basic argumentation components of warrant and backing; (2) conditioned argument and rebuttal: the text was divided into sections, each of which contained a certain argument; and (3) the level of collective argumentation in each section was determined according to the scope of the various components and aspects of the rebuttal.

The content validity was related to questions that were asked, as well as to the categories established by then researcher (the first author of this paper). The categories were based on the students' answers. The reliability between judges regarding adjusting the classification of students' answers to the categories was done later. The reliability correlation revealed an average of 87% among the judges. If there were answers for which the judges did not agree, then they met and discussed the problems until a consensus was reached. The level of argumentation used in the discourse was analyzed and classified on a scale ranging from 1 to 5, in accordance with Toulmin's argument model (Toulmin 1958), (Table 2).

The analysis was carried out through written reports. The arguments in them reflect the discourse that took place within the group and summarize it at two points in time: (1) The stage at which the hypothesis was formulated and (2) the stage at which the conclusions were written. The written text does not reflect the disagreements or rebuttals, only the content on which the group eventually agreed.

2.6 Interviews

The interviews were semi-structured. They provided information about how students perceived their role in the discourse that took place during the experiment or their contribution to it. Students were asked questions such as: "What is your opinion is the importance of collective discourse in general and specifically in a laboratory experiment?"; "How do you see your role in the experiment carried out in the laboratory?"; and "Your fellow group member says that the result of the experiment does not provide an answer to the research

Table 2 Discourse analysis of the level of argumentation based on (Erduran et al. 2004; Osborne et al. 2004; Katchevich et al. 2013)

Argument components	Abbreviations	Argument level
Claim	C	1
Claim + Data or Claim + scientific basis − Warrant	CD/CW	2
Claim + Data + Warrant or Claim + Data + Rebuttal or Claim + Warrant + Rebuttal	CDW/CDR/CWR	3
Claim + Data + Warrant + Backing	CDWB	4
Rebuttal that includes the Claim + Data + Warrant	CDWR	5

question; what do you think? The interviews were conducted in pairs and included above questions related to the students' opinions about learning science via experiments to make sure that they can use the laboratory as a vehicle for argumentation enhancement. The reliability for encoding the interview questions and the percentage of argument among the reviewers is 87%, that means higher level of argument of the reliability among the interviews, which demonstrates high level of argument between the judges and thus it, present high level of reliability in the interview tool of the research.

2.7 Research Procedure

In order to measure the evolution of the ability to construct arguments, three replicate confirmatory and three replicate open-ended experiments were conducted. The data were collected at three levels: Students' interviews, arguments that appear in the laboratory reports, and the number of arguments that developed during the discourse.

The time-on-task in the experiments was based on Katchevich's work (Katchevich et al. 2013). The duration of an open-ended inquiry experiment is significantly longer than a confirmatory one. The open-ended inquiry experiment is conducted over at least 6 lessons, whereas the confirmatory one is conducted over 4 lessons. Therefore, we checked the discourse during the confirmatory experiment from its beginning until its end—over the four lessons. Regarding the open-ended inquiry experiments, we referred to the discourse only when creating a hypothesis, analyzing the data, or drawing conclusions.

The experiments took place in two chemistry classes. At first, a confirmatory experiment (without sugar) was performed, followed by a classroom discussion under the researcher's guidance. The discussion concerned the formulation of the research question, the formulation of the research hypotheses, the execution of the experiment, analysis of the results, and writing down the conclusions. Then the open-ended experiment (with sugar) was carried out, again followed by a classroom discussion.

3 Results

The purpose of the study was to determine whether the chemistry laboratory activity served as an environment that significantly promotes learning and is suitable as a platform for improving students' argument construction skills.

The findings below relate to the following research questions:

1. What are the discourse features, according to students of science education while performing an open-ended laboratory experiment?
2. What are the discourse features, according to students of science education while performing the confirmatory laboratory experiment?

3.1 Comparison of the Number and the Level of Arguments and their Level in the Discourse Associated with Open-ended and Confirmatory Experiments

Data were collected on the observations of three different groups that performed the open-ended and confirmatory experiments: all three groups participated in both types of

experiments. Three groups of arguments were analyzed: (1) arguments that developed during the discourse, (2) arguments that appear in the laboratory reports, and (3) arguments that were mentioned in the interviews. The findings and analysis in this section refer to both types of experiments.

3.2 Analysis of the Discourse

A confirmatory experiment is one in which the students follow the teacher's instructions. Usually, the teacher's aim in this type of experiment is to confirm and reinforce the theoretical material learned in the classroom. An open-ended experiment is one in which the students decide what to study and how to proceed. In order to assess the level of the arguments, we chose a tool that refers to the various elements of an argument (see Tables 3–4, pages 8–9). This tool was chosen from among many assessment tools appearing in the literature, reviewed by Sampson and Clark (2008). This tool is aligned with the discourse style of the laboratory experiments and with Toulmin's model; it is based on other tools suggested in former studies (Erduran et al. 2004; Osborne et al. 2004; Simon and Johnson 2008). During the discourse, the students suggest different explanations for the various phenomena that they observed during the experimental procedure and then analyze the data and present arguments. The reliability of the coding of the argumentation discourse components was tested in two ways: encoding the components of the argumentation in 20% of the transcribed discourse and checking the reliability using three experts. The percentage of agreement between the experts ranged from 80 to 90%. For encoding in which the experts did not agree on, the judges discussed the issue until they reached a consensus. In addition, the authors repeated the encoding; after a while, the correlation between the early and late coding was 0.96.

From Tables 3, 4, and 5 we see that:

1. The number of arguments in the discourse during the open-ended experiment was 52, compared to 43 during the confirmatory experiment. Using the Kruskal-Wallis test, we found that this difference is significant: $Chi^2(1) = 4.52$, $p = 0.05$.
2. From the Chi^2 test for finding differences among argument frequencies at the different levels in the two types of experiments, no significant difference was found between the argument level and the experiment type: $Chi^2(4) = 8.74$, $p = NS$.

To conclude, the groups formulated more arguments in the discourse during the open-ended experiment than in the confirmatory one.

Table 3 Argument frequency in the discourse analysis of the confirmatory experiment groups according to the argument level

Experiment type	Argument level	Argument components	Group A	Group B	Group C	Total
Confirmatory experiment	1	C	6	7	10	23
	2	CD/CW	6	2	4	12
	3	CDW/CDR/CWR	2	3	0	5
	4	CDWB	0	0	1	1
	5	CDWR	1	1	0	2
	Total		15	13	15	43

Table 4 Argument frequency in discourse analysis of open-ended experiment groups according to the argument level

Experiment type	Argument level	Argument components	Group A	Group B	Group C	Total
Open-ended experiment	1	C	11	13	4	28
	2	CD/CW	2	2	2	6
	3	CDW/CDR/CWR	1	3	4	8
	4	CDWB	0	0	0	0
	5	CDWR	5	2	3	10
	Total		19	20	13	52

3.3 Analysis of the Laboratory Reports

The written arguments in the reports reflect the discourse that took place within the group. The report summarizes this discourse at two points in time: the hypothesis formulation stage and the summation stage.

The written reports do not contain the disagreements or rebuttals—only the contents on which there was a consensus within the group.

The reports were analyzed for argument frequency according to their level, the mean argument level, and the mean number of arguments in the reports on the open-ended and the confirmatory experiment.

From Tables 6, 7, and 8 we see that:

1. The number of arguments in the laboratory reports on the open-ended experiment was 37, compared to 25 in the laboratory reports on the confirmatory experiment. Using the Kruskal-Wallis test, we found that this difference is significant: $Chi^2(1) = 8.32$, $p = 0.005$.
2. From the Chi^2 test for finding differences among argument frequencies at the different levels in the two types of experiment, no significant difference was found between argument level and experiment type: $Chi^2(2) = 0.461$, $p = NS$.

Table 5 Frequency and percentage of arguments in the analysis of the discourse, according to the experiment type and argument level

	Experiment groups n (% of the experiment group)		Total
	Confirmatory experiment	Open-ended experiment	
Level 1	23	28	51
	(53.5%)	(53.8%)	(53.6%)
Level 2	12	6	18
	(27.9%)	(11.5%)	(18.9%)
Level 3	5	8	13
	(11.6%)	(8.4%)	(13.7%)
Level 4	1	0	1
	(2.3%)	(0%)	(1.1%)
Level 5	2	10	12
	(4.7%)	(19.2%)	(12.6%)
Total	43	52	95
	(45.3%)	(54.7%)	(100%)

3. Despite the lack of a significant difference, it should be noted that 56% of the arguments in the confirmatory experiment were at level 1, compared to 48.6% of level 1 arguments in the open-ended experiment (Table 8). Furthermore, 20% of the arguments in the confirmatory experiment were at level 2, in contrast to 27% of such arguments in the open-ended experiments, and 24% of the arguments in the confirmatory experiment were at level 3, whereas 24.3% of the arguments in the open-ended experiment were at this level.

The groups reported more arguments in their laboratory reports on the open-ended experiment than on the confirmatory one. In neither experiment type were there any arguments of levels 4 and 5.

3.4 Analysis of the Interviews

In the interviews, we attempted to determine whether students of science education perceive the different types of experiments as they are described in a survey of the literature (Domin 1999; Fradd et al. 2001; Herron 1971; Schwab 1962), with respect to the students' place in the different types and the distinct requirements of each type as well as with respect to defining the skills needed in them (Rosenberg 2007), as shown in Table 4. In order to accomplish this, interviews with students were conducted, posing a number of questions concerning basic-level experiments and confirmatory experiments versus open-ended experiments. Tables 9, 10, and 11 below present the data.

From Tables 9, 10, and 11 we see that:

1. The number of arguments in the interviews on the open-ended experiment was equal to the number of arguments in the interviews on the confirmatory experiment, 16 in each case.
2. From the Chi^2 test for finding the differences in argument frequencies at different levels in the two types of experiments, no significant difference was found between argument level and experiment type: $Chi^2(4) = 1.14$, $p = NS$.

Table 12 provides examples of arguments at various levels, taken from the data on discourse in the laboratory.

In this study, we have found that students who were involved in open-ended experiments posed more arguments in general and more high-level arguments in particular.

Table 6 Argument frequency in the laboratory reports of the confirmatory experiment groups according to the argument level

Experiment type	Argument level	Argument components	Group A	Group B	Group C	Total
Confirmatory experiment	1	C	6	4	4	14
	2	CD/CW	2	1	2	5
	3	CDW/CDR/CWR	3	1	2	6
	4	CDWB	0	0	0	0
	5	CDWR	0	0	0	0
	Total		11	6	8	25

Table 7 Argument frequency in the laboratory reports of the open-ended experiment groups according to the argument level

Experiment type	Argument level	Argument components	Group A	Group B	Group C	Total
Open-ended experiment	1	C	7	4	7	18
	2	CD/CW	4	3	3	10
	3	CDW/CDR/CWR	2	3	4	9
	4	CDWB	0	0	0	0
	5	CDWR	0	0	0	0
	Total		13	10	14	37

4 Discussion

The main objective of this research study was to examine scientific activity in a laboratory setting as a significant learning-promoting environment and as a means for advancing argument-construction skills. These aspects were examined through the discourse that took place during the laboratory activity and in the reports written after each experiment, as well as in the classroom discussion on laboratory-related issues that occurred after each experiment. The data show that the number of arguments in the open-ended experiment was greater than in the confirmatory experiment, and that the arguments in the former case were at a higher level than in the latter case. We suppose that students how exposed to open-ended (inquiry) experiment demonstrate significant more than those how exposed to confirmatory experiments.

This led to the assumption that the laboratory can serve as an environment that encourages an argument-constructing discourse with no need for intervention, due to the unique features of this environment: working in small groups that allow students to conduct a discourse, and choosing a practical topic about the possibility of conducting a discourse, but at the level of understanding the stages of a scientific experiment and at the level of implementation, obtaining results, and discussing them.

Our study found that when students obtain unexpected results in an experiment that they plan, the ensuing discourse contains more arguments as well as rebuttals, which arouse a cognitive conflict in the learners; this drives them to reexamine what they know, in order to discover why their previous knowledge did not constitute a sufficiently good basis to explain

Table 8 Frequency and percentage of arguments in the laboratory reports according to the experiment type and argument level

	Experiment groups n (% of the experiment group)		Total
	Confirmatory experiment	Open-ended experiment	
Level 1	14	18	32
	(56%)	(48.6%)	(51.6%)
Level 2	5	10	15
	(20%)	(27%)	(24.2%)
Level 3	6	9	15
	(24%)	(24.3%)	(24.2%)
Total	25	37	62
	(40.3%)	(59.7%)	(100%)

Table 9 Argument frequency in the interviews of members of the confirmatory experiment groups according to the argument level

Experiment type	Argument level	Argument components	Group A	Group B	Group C	Total
Confirmatory experiment	1	C	1	1	1	3
	2	CD/CW	1	2	1	4
	3	CDW/CDR/CWR	0	1	1	2
	4	CDWB	2	2	1	5
	5	CDWR	1	0	1	2
	Total		5	6	5	16

the results, leading them to expand their knowledge or to offer explanations with a different scientific basis, one which they had not considered previously or had not been aware of.

In order to confirm our claim concerning the difference in the number and level of arguments in open-ended versus confirmatory experiments, we compared the arguments in both types of experiments conducted by the same group. The findings (Tables 5, 8, and 11) provide support for our claim that the difference in the number and level of the arguments lies in the task itself, and not in the many other possible intervening factors that could have affected the discourse.

Note that despite the significant difference in the number and level of arguments found in the discourse conducted by our experimental groups during the two types of experiments, no significant difference in the level of the arguments was found in the laboratory reports on the two types. It is possible that in the wake of formulating their conclusions in their reports on the open-ended experiment, students realized what was expected of them when writing their conclusions, and therefore they applied this skill in the case of the confirmatory experiment as well.

Our findings are consistent with those of previous studies (Erduran and Kaya 2018; Tien and Stacy 1996) that concluded that students who learned using the guided inquiry method were better able to assess the evidence of the research to which they had been exposed, and that their explanations were more substantial, and their arguments were accompanied by better explanations. Furthermore, the findings lend support to the approach adopted by Osborne et al. (2004), who proposed strategies that include activities for developing argumentation skills, such as exposing students to different claims concerning a given scientific topic, then asking students to give supporting arguments in favor of both theories, one of them, or neither. In view of the fact that guided inquiry was found to be an appropriate teaching strategy (Wilson et al. 2010), the importance of laboratory work, as stated in the past (Hodson 1993), reinforces the claim that well-planned and well-executed inquiry laboratory experiments can promote

Table 10 Argument frequency in the interviews of members of the open-ended experiment groups according to the argument level

Experiment type	Argument level	Argument components	Group A	Group B	Group C	Total
Open-ended experiment	1	C	1	1	1	3
	2	CD/CW	1	1	1	3
	3	CDW/CDR/CWR	0	0	1	1
	4	CDWB	2	2	1	5
	5	CDWR	1	1	2	4
	Total		5	5	6	16

Table 11 Frequency and percentage of arguments in interviews according to the experiment type and the argument level

| | Experiment groups n (% of the experiment group) | | Total |
	Confirmatory experiment	Open-ended experiment	
Level 1	3	3	6
	(18.8%)	(18.8%)	(18.8%)
Level 2	4	3	7
	(25%)	(18.8%)	(21.9%)
Level 3	2	1	3
	(12.5%)	(6.3%)	(9.4%)
Level 4	5	5	10
	(31.3%)	(31.3%)	(31.3%)
Level 5	2	4	6
	(12.5%)	(25%)	(12.8%)
Total	16	16	32
	(50%)	(50%)	(100%)

learning, understanding of concepts, and understanding the nature of science among students (Hofstein and Lunetta 1982; Tobin 1990).

5 Conclusions

The present study aimed to investigate classroom discourse and the argument construction process among students of science education in the context of conducting an experiment, both in the discourse that takes place during the experiment itself and in the subsequent classroom discussion on topics that arose during the experiment. The data gleaned from the research tools were analyzed and the argument levels found in them were classified according to a scale ranging from 1 to 5 derived from Erduran et al. (2004).

We found that for the research groups that we observed, the laboratory was able to serve as a platform for argument construction without any intervention, owing to this learning environment's unique features: working in small groups, which made it possible to develop a discourse and an environment that provided students with time and a platform (Lazarowitz and Tamir 1994).

We found that when students obtain unexpected results in an experiment that they planned, the developing discourse contains more arguments as well as rebuttals.

In order to reinforce our claim concerning the difference in the number and level of arguments, depending on whether the experiment was open-ended or confirmatory, we compared the arguments presented by the same group in both types of experiments. The findings lent support to our claim.

Note that despite the significant difference in the number and level of arguments found in the discourse conducted by our experimental groups, students became aware of the need to justify their decisions with scientific evidence. It was concluded that introducing argumentation about experimental issues to students in a school can improve their argumentation skills (Dawson and Carson 2018). No significant difference in the level of the arguments was found in the laboratory reports by the two types of groups. It is possible

Table 12 Examples of arguments at different levels, from data on discourse in the laboratory

	Argument level	Code	Argument component
Rami: We can see that the dough in the first glass is full of bubbles. Probably it is carbon dioxide	1	C	Claim
Hiba: The yeast is the main factor in the rise of the dough. As we saw in the first glass, "dough with yeast" the dough rises and floats on the water, whereas the ball of "dough without yeast" remains sunken at the bottom of the glass	2	CD CW	Claim + Data, or Claim + Scientific basis – Warrant
Hayya: There are bubbles (holes) on the dough. From this, we can conclude that yeast consists of unicellular microorganisms of the fungus group that are carrying out respiration. In this process, carbon dioxide is created and is trapped inside the dough	3	CDW CDR CWR	Claim + Data + Warrant, or Claim + Data + Rebuttal, or Claim + Warrant + Rebuttal
Ghada: Why does the dough with yeast in the first glass rise and float on the water, whereas the ball of dough without yeast sinks to the bottom?	4	CDWB	Claim + Data + Warrant + Backing
Hala: Because the yeast, consisting of unicellular microorganisms of the fungus group, emits carbon dioxide during cell respiration. When food is available and the temperature is appropriate for yeast, they emit carbon dioxide. As a result, the dough expands and its specific weight drops, so that it floats on water			
Fida: Maybe we put sugar in both glasses by mistake	5	CDWR	Rebuttal that includes: Claim + Data + Warrant
Suha: We did not. Only in glass number 1			
Rana: Maybe we made a mistake in the research procedure we planned. Should we review the procedure?			
Suha: Did the sugar cause an increase in the mass of the dough?			
Fida: What can we conclude from this?			
Suha: We can conclude that in both glasses there are equal quantities of flour and yeast, and that the only difference is the sugar			
Fida: So what made the difference?			
Suha: The yeast, consisting of unicellular microorganisms of the fungus group, emits carbon dioxide during cell respiration. In this case, the food is sugar with flour; thus the volume of dough grows and its specific gravity decreases, so that it floats on water. In comparison to the first glass with the added sugar, it apparently interfered with cell respiration and caused the dough not to rise, so that it sank in the water			

that in the wake of formulating their conclusions in their reports on the open-ended experiment, students realized what was expected of them when writing their conclusions, and that therefore they applied this skill in the case of the confirmatory experiment as well. In fact, despite the fact that their discourse was relatively poor in high-level arguments,

when they wrote down their conclusions on the confirmatory experiment, some of them wrote conclusions similar to those that they had written in the case of the open-ended experiment.

5.1 Practical Implications for Science Education

The study's findings offer teachers a way to teach argumentation skills through discourse and therefore make it possible to begin to bring about a change in teachers' perception of science teaching and in students' perception of science.

1. We recommend constructing a set of lectures on acquiring argumentation construction skills. If teachers do not master these skills, they will be unable to model them for their students, and the students will be unlikely to acquire them.
2. We recommend that teachers explicitly stress the importance of group discussions during the laboratory experiments and during classroom activity, based on learning theories.

5.2 Future Investigations

A large number of arguments do not necessarily mean that the arguments have scientific quality (Puvirajah 2007). Therefore, it is also expected to have argumentation schemes that can be classified as scientific and involve justified conclusions, since the laboratory work will provide empirical evidence for the construction and evaluation of arguments (Ozdem et al. 2013). For future investigations and for analyzing students' arguments for their quality and scientific credibility, we suggest using, in addition to Toulmin's model, another model such as a model called the Structure of Arguments and Scientific Credibility Model (SASC) (Puvirajah 2007) in order to investigate how pre-service teachers will develop their higher-order thinking and scientific inquiry habits so that they become proficient in all the components of scientific inquiry and in formulating high-quality and highly credible arguments.

Compliance with Ethical Standards

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

Ethical Standards and Issues: a Declaration   The research described in this paper was conducted among second year college students studying biology (pre-service). We asked the students for their consent to participate in this study that included both recording their discussions (in-group) as well as interviews. We promised the students to use the information for only research purposes and not for assessing them in terms of their knowledge and attitudes. The results were not used for grading purposes.

References

Bybee, R. W. (2000). Teaching science as inquiry. In J. Minstrel & E. H. Van-Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 20–46). Washington DC: AAAS.
Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: a practical guide. *The Journal of the Learning Sciences, 6*(3), 271–315.

Dawson, V., & Carson, K. (2018). Introducing argumentation about climate change socio scientific issues in a disadvantaged school. *Research in Science Education*. https://doi.org/10.1007/s11165-018-9715-x.

Erduran, S. (2018). Toulmin's argument pattern as a "horizon of possibilities" in the study of argumentation in science education. *Cult Stud of Sci Educ*. https://doi.org/10.1007/s11422-017-9847-8.

Erduran, S., & Kaya, E. (2018). Drawing nature of science in pre-service science teacher education: epistemic insight through visual representations. *Research in Science Education, 48*, 1133–1149.

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., 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*(5), 1–12.

Gott, R., & Duggan, S. (2007). A framework for practical work in science and scientific literacy through argumentation. *Research in Science and Technological Education, 25*(3), 271–291.

Hobson, A., Ashby, P., Malderez, A., & Tomlinson, P. D. (2009). Mentoring beginning teachers: what we know and what we don't. *Teaching and Teacher Education, 25*(1), 207–216.

Hodson, D. (1993). Re-thinking old ways: Toward a more critical approach to practical work in school science. *Studies in Science Education, 22*(1), 85–142.

Hofstein, A., & Kind, P. (2012). Learning in and from science laboratories. In B. Fraser, K. Tobin, & K. McRobbie (Eds.), *Second handbook of research in science teaching* (pp. 189–208). Dordrecht Nl: Springer.

Hofstein, A., & Lunetta, V. N. (1982). The role of the laboratory in science teaching: neglected aspects of research. *Review of Educational Resaerch, 52*(2), 201–217.

Hofstein, A., Shore, R., & Kipnis, M. (2004). Providing high school chemistry students with opportunities to develop learning skills in an inquiry-type laboratory: a case study. *International Journal of Science Education, 26*, 47–62.

Hofstein, A., Kipnis, M., & Kind, P. (2008). Learning in and from science laboratories: enhancing students' meta-cognition and argumentation skills. In C. L. Petroselli (Ed.), *Science education issues and development* (pp. 59–94). New York: Nova Science.

Jaspers, W. M., Meijer, P. C., Prins, F., & Wubbels, T. (2014). Mentor teachers: their possibilities and challenges as mentor and teacher. *Teaching and Teacher Education, 44*(11), 106–116.

Katchevich, D., Mamlok-Naaman, R., & Hofstein, A. (2013). Argumentation in the chemistry laboratory: inquiry and confirmatory experiments. *Research in Science Education, 43*(1), 317–345.

Katchevich, D., Mamlok-Naaman, R., & Hofstein, A. (2014). The characteristics of open-ended inquiry-type chemistry experiments that enable argumentative discourse. *Journal of Education, 2*(2), 74–99.

Kipnis, M., & Hofstein, A. (2008). The inquiry laboratory as a source for development of metacognitive skills. *International Journal of Science and Mathematics Education, 6*, 601–627.

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

Lazarowitz, R., & Tamir, P. (1994). Research on using laboratory instruction in science. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp. 94–128). New York: Macmillian.

McDonald, C. V. (2014). Preservice primary teachers' written arguments in a socioscientific argumentation task. *Electronic Journal of Science Education, 18*(7), 1–20.

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council. (2013). Next generation science standards: for states, by states. Washington, DC: The National Academies Press. https://doi.org/10.17226/18290.

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

Ozdem, Y., Ertepinar, H., Cakiroglu, J., & Erduran, S. (2013). The nature of pre-service science teachers' argumentation in inquiry-oriented laboratory context. *International Journal of Science Education, 35*(15), 2559–2586.

Puvirajah, A. (2007). Exploring the quality and credibility of students' argumentation: teacher facilitated technology embedded scientific inquiry) . *Dissertation Abstracts International, 68*(11) (UMI no. 3289408).

Sampson, V., & Clark, B. (2008). Assessment of ways students generate arguments in science education: current perspectives and recommendations for future directions. *Science Education, 92*, 447–472.

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

Simon, S., & Johnson, S. (2008). Professional learning portfolios for argumentation in school science. *International Journal of Science Education, 30*, 669–688.

Therrien, J., Taylor, J. C., Watt, S., & Kaldenberg, E. (2014). Science instruction for students with emotional and behavioral disorders. *Remedial and Special Education, 35*(1), 15–27. https://doi.org/10.1177/0741932513503557.

Tien, L. T., & Stacy, M. (1996). The effects of instruction on undergraduate students' inquiry skills. Paper presented at the meeting of the American Educational Research Association, New York.

Tobin, K. (1990). Research on science laboratory activities: in pursuit of better questions and answers to improve learning. *School Science and Mathematics, 90*(5), 403–418.

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

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the Development of Children, 23*(3), 34–34.

Wilson, C. D., Taylor, J. A., Kowalski, S. M, & Carlson, J. (2010). The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. *Journal of Research in Science Teaching, 47,* 276–301.