9 DEVELOPING THE ABILITY TO CRITIQUE IN THE COURSE OF INQUIRY-ORIENTED PROGRAMS IN BIOLOGY

Tom Bielik and Anat Yarden
Department of Science Teaching, Weizmann Institute of Science, Rehovot, Israel
tom.bielik@weizmann.ac.il, anat.yarden@weizmann.ac.il

Abstract

Authentic scientific practices are designed to facilitate students' understanding of how scientific knowledge develops, including the ability to critique, which constitutes an important part of scientific inquiry. Students should be able to identify potential weaknesses and flaws in scientific claims, articulate the merits and limitations of peer views and read media reports in a critical manner. Even though the importance of incorporating critique in science education classrooms is well accepted and emphasized by the science education research community, much debate still remains regarding how this practice should be taught. We set out to explore the contribution of an inquiry-oriented program for high-school students which emphasizes critiquing. Pre- and post-questionnaires were administered to students participating in an inquiry-oriented program (Bio-Tech), and to students who were not participating in the program. Students of both groups tended to be more in agreement with an arguable claim presented to them in the post-questionnaires compared to the pre-questionnaires. However, the Bio-Tech students tended to use more arguments and focused more on the experimental process described to them than the Control group students. These results indicate that students can develop some critiquing abilities in the context of an inquiry-oriented program in biology.

Keywords: Inquiry; Critique; Scientific practice; Authenticity; Argumentation

1. Introduction

Most recent policy documents present the ongoing call for successful implementation of authentic scientific practices in science classrooms (European Commission, 2007; National Research Council [NRC], 2000, 2012). The ability to practice inquiry requires that students not only learn the traditional process skills, but also combine them with scientific knowledge, reasoning and the ability to critique. Authentic scientific practices include not only skills but also specific knowledge required for investigating and building models and theories about the natural world (National Research Council [NRC], 2012). Much emphasis is directed to the social and cognitive aspects of the scientific process: the communication, argumentation and model-generating practices. Authentic scientific practices are designed to facilitate students' understanding of how scientific knowledge develops, and of 'scientific habits-of-mind' and engagement in scientific inquiry (National Research Council [NRC], 2012; Osborne, 2010).

The ability to critique is generally defined as "reasonable reflective thinking that is focused on deciding what to believe or do" (Ennis, 1987). The ability to critique makes up an important part of scientific inquiry and consists of overlapping skills and abilities, such as testing hypotheses, designing experiments and drawing conclusions from results (Berland & Reiser, 2009; Ford, 2008). Students should be able to identify possible weaknesses and flaws in scientific claims, articulate the merits and limitations of peer views and read media reports in a critical manner (National Research Council [NRC], 2012). The ability to critique is crucial for productive participation in scientific practice and discourse (National Research Council [NRC], 2007). Berland and Reiser (2011) considered critiquing to be a key part of the goals of sense-making and persuasion in scientific argumentation.

Critiquing is strongly connected to the practice of argumentation, which is one of the central goals of science education and the focus of several recent articles and policy documents (Berland & McNeill, 2010; National Research Council [NRC], 2007, 2012; Osborne, 2010). Argumentation is connected to other scientific skills and abilities, such as reasoning, critical and logical thinking, language skills, communication and justification. An argument is defined as an assertion or conclusion with justification, reasons and support (Osborne et al., 2004). Ford (2008) reported that scientists are more likely to have less confidence in a given scientific claim and that their critique mostly concerns the methods used to collect the data and the analysis and evaluation of the results. Non-scientists, on the other hand, are more likely to accept the given scientific claims and relate their reasoning arguments mostly to their personal experiences. In a more recent work, Ford (2012) claimed that constructing and critiquing arguments are fundamental parts of scientific sense-making during engagement in scientific discourse.

Even though the importance of incorporating critique in science education classrooms is well accepted and emphasized by the science education research community, much debate still remains on how this practice should be taught. Osborne (2010) argued that students in contemporary classrooms lack the opportunity to develop and master their abilities to reason out and critique scientific claims. It was suggested that students rarely have opportunities to be engaged in critiquing and in scientific argumentation because traditional approaches to science instruction do not promote or support student engagement in scientific argumentation (Sampson & Clark, 2011). Others indicated that students, in general, lack the abilities to construct and present arguments and are poor at addressing different points of view regarding learned scientific issues. It was claimed that more activities are needed to develop these abilities in the classroom, mainly by restructuring current science lessons (Berland & Reiser, 2011; Driver et al., 2000).

Appropriate means of incorporating critique in science classrooms remain to be clarified and explored. There is a need to characterize the development of critiquing ability among students in science classrooms and to explore possible activities which can engage students in this activity. Here we suggest that inquiry-oriented scientific programs are adequate as a platform for developing students' ability to critique, providing the appropriate support to teachers and the scientific environment.

In this study, we explore the contribution of an inquiry-oriented program for high-school students which emphasizes critique. Our aim is to characterize and evaluate possible changes in students' arguments in response to an arguable claim made by a hypothetical student, focusing on their tendency to agree or disagree with the claim, the number of arguments they use in their answer in response to the claim, the categories of arguments they use and their qualitative characteristics. Our research question is whether participation in an inquiry-oriented program improves high-school biotechnology majors' ability to critique. In order to answer this question, we set to examine whether students who participate in the inquiry-oriented program tend to be in agreement with peer claims, do they use more arguments in response to peer claims and whether they focus their arguments more on the experimental process, methods or chain of inferences.

2. Research design and method

This research was designed to evaluate and characterize possible changes in students' ability to critique following their participation in an inquiry-oriented program in biology termed Bio-Tech program. Pre- and post-questionnaires were administered to 11th-grade biotechnology majors who were either participating or not participating in the Bio-Tech program. The questionnaires included a scientific article and a deliberately arguable hypothetical student's claim.

2.1 Research context

The Bio-Tech program at the Weizmann Institute of Science (hereon referred to as 'the Bio-Tech program') is an optional part (1 credit out of a total of 5 credits) of the Israeli matriculation examinations for biotechnology majors during the 11th grade (Israeli Ministry of Education, 2005). It is based on a visit to a biotechnology laboratory in an industrial or academic facility. The Weizmann Institute began supporting the Bio-Tech program in 2009 and the current research was carried out during the 2011/12 academic year. The Bio-Tech program design originates from the Teacher-Led Outreach Laboratory (TLOL) program that is practiced at the Weizmann Institute (Stolarsky Ben-Nun & Yarden, 2009).

The Bio-Tech program is unique and innovative in the following aspects: the inquiry-based approach allows students to practice high levels of open inquiry, a co-teaching approach is implemented (teaching is performed by the class teacher, a research scientist, and a science educator), and the topic of inquiry is learned using the Adapted Primary Literature (APL) approach with an adapted scientific article. This allows the students to learn up-to-date scientific concepts, practice technologically advanced methods and tools and experience a firsthand encounter with authentic science (Yarden et al., 2001).

The investigated biological systems range from the molecular and genetic level, including proteins and organelles, to the living organism level of bacteria, fungi, yeast, and tissue-cultured cells. Currently, six research groups from the Weizmann Institute and from the Robert H. Smith Faculty of Agriculture, Food and Environment of the Hebrew University are taking part in the Bio-Tech program. The techniques used in this program range from simple observational methods (such as bacterial colony growth on plates, color changes in medium, microscope observation) to the use of highly advanced tools and equipment (such as spectrophotometer, PCR, fluorescence microscope). The protocols are specially designed and adapted to fit the students' cognitive abilities and the time constraints of the program.

The Bio-Tech program is carried out during an entire academic school year. It is comprised of learning the background knowledge using an APL article, a preliminary visit to the research institute where students visit the particular laboratory related to their specific project and perform a series of short experiments in which they acquire key concepts and techniques related to the specific inquiry project, formulating the research questions and planning the main experiments in dyads back in the classroom, performing the experiment in a two days main visit to the research institute and analyze their findings and prepare their research portfolio in a 2-5 months long process back in school with the assistance of the teacher. The final grade of each student is determined based on an oral examination which takes place around the end of the school year, conducted by an external examiner (a biotechnology teacher from another school) and the class teacher.

In the Bio-Tech program, much emphasis is explicitly directed to developing the students' ability to critique and articulate their own knowledge and claims. At the beginning of the program, when students study the APL paper, they are engaged in classroom discussions, led by the teacher, in which they are confronted with the scientific knowledge together with the reasons for using the specific scientific methods and tools. They are expected to understand the scientific content and process by the time they arrive at the research laboratory for their preliminary visit. When formulating their research question and planning the experiment, students are actively engaged in communicating with their peers and their teacher. They learn how to defend and explain their research question and are expected to master all stages of the planned experimental process. During their discussions with the teacher, the scientist and the science educator, students are frequently required to justify what they do, to demonstrate their understanding of the research and to explain their results and analysis. Although this process is long and sometimes frustrating for the students, the class instructors are well trained and experienced in providing adequate support and guidance for the students. In the final part of the program, students write a scientific report in the form of a research article, which is a major part of the research portfolio. In the oral examination, the student is expected to defend his/her work and justify its conclusions, as well as present both content and procedural understanding. Taken together, during the Bio-Tech program, students are given numerous opportunities to develop their ability to critique.

Some specific activities, designed for developing the Bio-Tech students' peer-critique and critique abilities, were incorporated into the program. For example, when dyads of students are working on formulating their research question and hypothesis, they are requested to choose among several research questions that they generate and to present the chosen question to another dyad. The other dyad is expected to review and critique the question according to the teachers' instructions. Following this activity, the original dyad receives their peer-reviewed question and asked to relate and consider the comments and to formulate their final research question to be presented to the teacher for further review and approval

2.2 Population

The research population was comprised of 11^{th}-grade biotechnology majors (16-17 years old). Four classes participating in the Bio-Tech program (the Bio-Tech group) and four classes not participating in this or in any other inquiry-oriented program (the Control group) were chosen. In total, 73 students from the Bio-Tech group and 58 students from the Control group filled in both pre- and post-questionnaires.

2.3 Tools

Pre- and post-questionnaires were designed to investigate students' identification of authentic scientific practices in a popular scientific article ('Alarm sounds over toxic teething rings', The New Scientist, July 14, 1997). After reading the article, students were given an arguable statement from a hypothetical student claiming a specific conclusion regarding the article ("This article *proves* that teething rings hurt babies" emphasis in original). This method was based on the previously published work of Ford (2012).

The article discusses the biological health issue of toxins released from babies' teething rings and its implications on their health. In the article, an experiment that was performed is presented, describing the methods and obtained results. After reading the article, students were asked to answer several open-ended questions designed to evaluate their understanding of the inquiry process presented in the article and to explore their question-asking practice. In one of the questions, students were given the hypothetical student's arguable claim (see above) and asked if they agree or disagree with the claim and why. The claim was deliberately arguable, and students were provoked to critique it from various aspects, such as the certainty and confidence level of the claim, the lack of evidence to support this claim and the flaws in the chain of inferences. The pre-questionnaires were administered at the beginning of the school year, before the selected classes had engaged in the Bio-Tech program. The post-questionnaires were administered at around the same time as the oral exam for the Bio-Tech students at the end of the school year.

2.4 Analysis

Only questionnaires of students who answered both the pre- and post-questionnaires were taken for analysis. Each answer was classified according to the students' agreement or disagreement with the arguable claim and the arguments they used were analyzed and categorized. Initial categories, depicted in a bottom-up process by the first author, were reviewed and validated by the second author and two other science education researchers. The classification of arguments to the different categories was unanimous in over 80% of the cases. The non-agreeable categories and arguments were further discussed until an agreement between the validators was reached regarding the classification of the arguments.

Students' answers were statistically analyzed using Statistical Analysis System (SAS) program for both descriptive statistics and comparing frequencies (Chi-square comparing). Results were statistically analyzed using the Wilcoxon signed-rank test for significant differences (Wilcoxon, 1945) and McNemar's test (Siegel & Castellan, 1988). Agreement or disagreement with the arguable claim was calculated as the percentage of students from the total number of students who answered the questionnaire in each group.

To categorize students' arguments, in-depth analysis of their answers was performed. Students' answers were classified into three main categories: (1) arguments regarding the different stages of the experiment described in the article (the 'described experiment' category), excluding arguments relating to the connection between the experimental results and the conclusions, which were classified in the second category, (2) arguments concerning the 'chain of inferences', namely the arguments made by the hypothetical student that connect the experimental results and the conclusions, and (3) arguments focusing on other issues presented in the article. The first category of arguments regarding the experiment described in the article was further split into the following three subcategories: (1) general arguments, (2) arguments focusing on the experimental process and protocol, and (3) arguments concerning the experimental conditions. The categories, subcategories and examples are detailed below (Table 1). Students' arguments in response to the arguable claim were qualitatively classified into the above categories and quantitatively analyzed.

Table 1. Categories of students' arguments regarding the hypothetical student's arguable claim

Category	Subcategories	Examples
1. Described experiment	A. General	*"I agree with the student because the article presents the results of a scientific experiment that proves that teething rings release a toxic substance that damages the baby." (Bio-Tech, #21)*
	B. Experimental process	"*I disagree with the student's opinion because the experiment was only performed once with no control and no repeats.*" (Bio-Tech, #5)
	C. Experimental conditions	*"The conditions under which the experiment was performed do not match the conditions under which babies use the teething rings." (Control, #23)*
2. Chain of inferences		"*I agree with the claim because we really see in the experiment that the rings release huge amounts of dangerous poisons.*" (Control, #5)
3. Other issues in the article		"*I disagree with the student…The article mentions that these substances may cause cancer, but it is not certain.*" (Control, #28)

3. Results

3.1 Students' responses to the arguable claim

To examine the possible changes in students' tendency to critique an arguable claim made by a hypothetical student following their participation in the Bio-Tech program, students' answers to the pre- and post-questionnaires were analyzed and compared to those of the Control group who did not participate in any inquiry-oriented program (Figure 1). No significant differences were found between the Bio-Tech and the Control groups in the pre-questionnaire regarding the percentage of students agreeing or disagreeing with the arguable claim ($p>0.05$).

A decrease in the percentage of students who disagreed with the arguable claim was observed in both the Bio-Tech and Control groups (from 64% to 49% and from 69% to 53%, respectively). This decrease was found to be statistically significant in both groups according to McNemar's test (Bio-Tech chi-square=4.17, $p<0.05$; Control chi-square=4.26, $p<0.05$). This decrease was accompanied by an increase in the percentage of students who agreed with the arguable claim in both groups (Biotech from 30% to 49%, chi-square=7, $p<0.01$; Control from 27% to 40%, chi-square=3.26, $p=0.07$).

A more detailed analysis of the shift from disagreement with the arguable claim in the pre-questionnaire to agreement in the post-questionnaire showed that a high percentage of both the Bio-Tech and Control group students shifted from disagreement to agreement (26% and 17%, respectively) with no significant differences between the two groups.

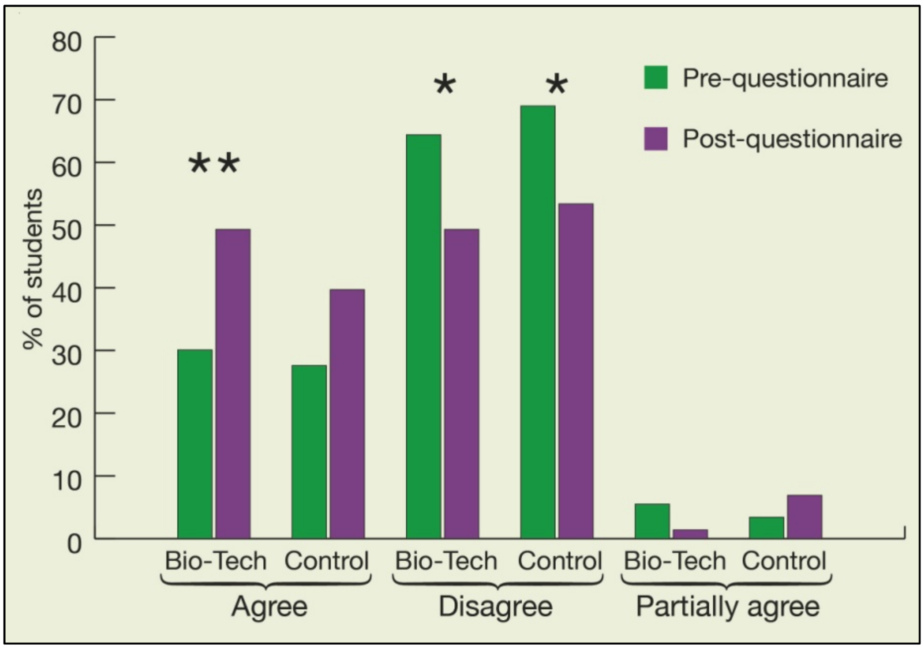

Figure 1. Comparison of students' positions toward the arguable claim in pre- and post-questionnaires (Bio-Tech n=73, Control n=58, *p<0.05, **p<0.01).

An example of students' tendency to shift from disagreement to agreement with the arguable claim, seen in both the Bio-Tech and Control groups, can be found in the analysis of one of the student's answers. This Bio-Tech group student (#55) disagreed with the arguable claim in his pre-questionnaire answer, using arguments related to the chain of inferences (*"I disagree with the student since this article didn't prove that all of the teething rings are dangerous for babies. It proved that there are specific kinds of teething rings that release phthalates and are dangerous for use, but that there are other teething rings which are not considered dangerous."*). In the post-questionnaire, the same student changed his opinion, agreeing with the claim and using arguments related to the experiment described in the article (*"I agree with the student since after establishing the hypothesis, the researchers performed the experiment in order to prove their hypothesis and with the experiment they proved that teething rings are dangerous for babies because of the phthalates that are released from them").*

In summary, students of both the Bio-Tech group and the Control group tended to be more in agreement with the arguable claim in the post-questionnaire, indicating that participation in the Bio-Tech program did not make the students more opposed to or less likely to agree with a peer's claim.

3.2 The number of arguments used by the students

We then explored possible changes in the number of arguments used by students in their answers following participation in the Bio-Tech program. We assumed that an increase in the average number of arguments might indicate a possible change in the students' ability to critique. However, no significant differences were found in the average number of arguments used by the Bio-Tech group students in the pre- and post-questionnaires (1.69 and 1.67, respectively, Figure 2). On the other hand, a statistically significant ($p<0.05$) decrease in the average number of arguments was found among students of the Control group (1.84 and 1.39, respectively, Figure 2). This indicates that the ability to use arguments was retained by the Bio-Tech students, while this ability showed a regression among students who did not participate in the inquiry-oriented program.

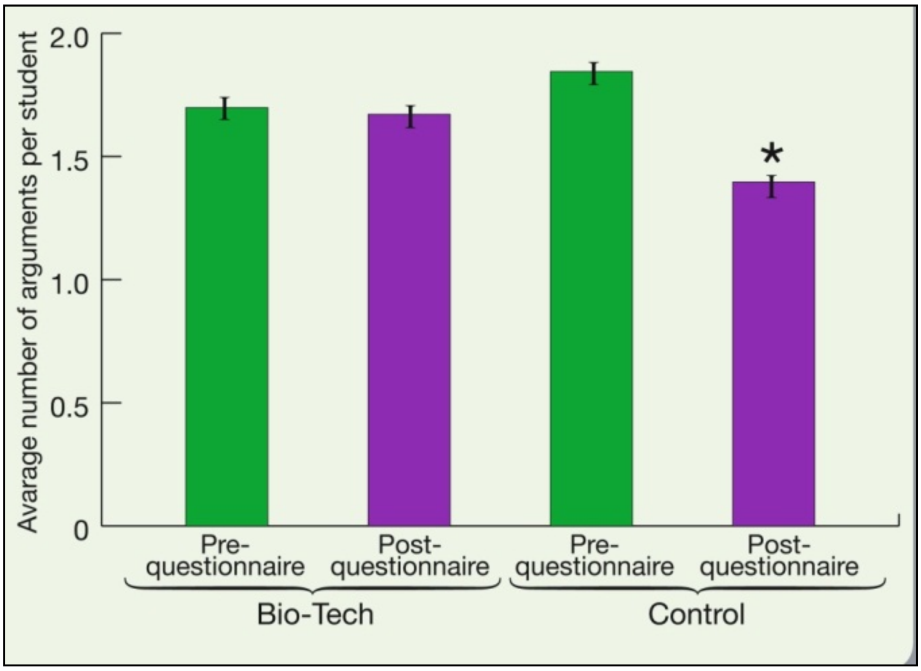

Figure 2. Average number of student arguments in pre- and post-questionnaires (Bio-Tech n=73, Control n=58, *$p<0.05$).

An example of the decreased average number of arguments in the answers of Control group students is presented in the following quote. This student (#55) from the Control group, who did not participate in the Bio-Tech program, disagreed with the arguable claim in the pre-questionnaire, using three arguments from the category of 'chain of inferences' ("*I disagree with the student, since the experiment in the article was performed on only 11 types of teething rings and this is not enough to determine and generalize that all teething rings are dangerous. There may be other companies that are not using this substance*"). In her post-questionnaire, however, this student agreed with the arguable claim and used only one argument in her answer ("*I agree. The article shows an experiment that proves that the teething rings are dangerous*").

3.3 In-depth analysis of students' arguments

To further explore the students' arguments and understand the possible changes in their arguments before and after the intervention, an in-depth investigation of the type of arguments used by the students was carried out. Students' answers were classified into categories and subcategories, as detailed in the methods section.

Classification of the students' arguments revealed that most of them, in both the Bio-Tech and Control groups, focused on the chain of inferences in both pre- and post-questionnaires

(Figure 3). There was a significantly (p<0.005) higher percentage of arguments related to the experiment described in the article in the pre-questionnaires compared to the post-questionnaires among the Bio-Tech group (from 10.6% to 25.6%), while no statistically significant change was observed among the Control group students according to Wilcoxon test.

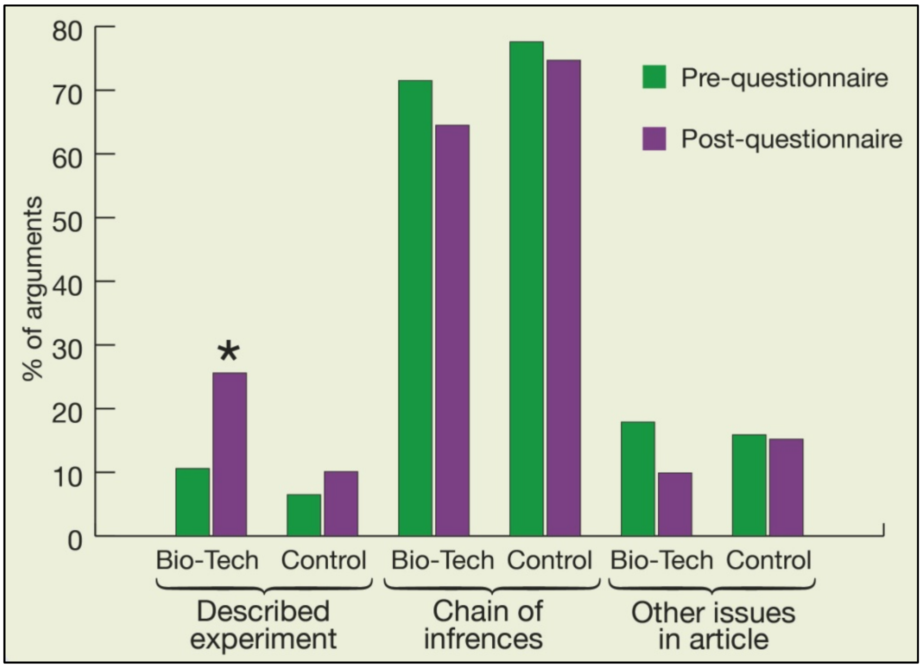

Figure 3. Comparison of students' argument types in pre- and post-questionnaires (Bio-Tech pre n=124, Bio-Tech post n=122, Control pre n=107, Control post n=81, *p<0.005).

An example of the increased tendency of Bio-Tech students to use arguments relating to the experiment described in the article is presented here. One of the students (#27) from the Bio-Tech group wrote an answer in the pre-questionnaire which included an argument from the category of other issues in the article, specifically arguments concerning the health issues of babies who use teething rings (*"I don't agree with the student. It was not experimentally examined or written in the article if phthalates are dangerous for babies or how they affect them. Maybe babies have immunity to phthalates? They didn't examine the activity of the baby who uses the teething rings compared to a baby who does not, therefore you can't know if the teething rings are dangerous."*). In the post-questionnaire, however, the same student still disagreed with the arguable claim but used arguments from the category of the chain of inferences (*"I disagree. The third ring released only 9 mg of phthalates and this amount is small and harmless")*. In addition, he used an argument from the category of the described experiment ("*They need to repeat the experiment to validate the results, examine all kinds of rings and only then determine which rings are dangerous").*

A closer examination of the total number of arguments used by the Bio-Tech students that are related to the category of the described experiment (Figure 4) revealed an increase in the post-questionnaires in all three subcategories: general issues of the experiment (from 2 arguments in the pre-questionnaire to 7 in the post-questionnaire), the experimental process (from 8 arguments in the pre-questionnaire to 17 in the post-questionnaire) and the experimental conditions (from 3 arguments in the pre-questionnaire to 7 in the post-questionnaire). This indicates improvement in the Bio-Tech students' ability to critique all aspects of the experiment presented to them.

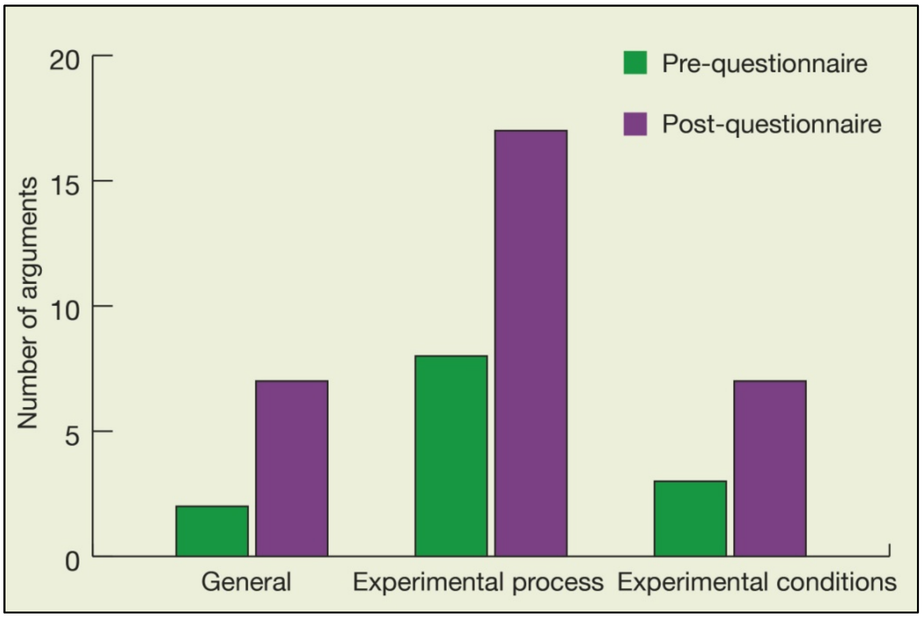

Figure 4. Number of Bio-Tech students' arguments related to the experiment described in the article (Bio-Tech, n=73).

Altogether, the results show that even though the overall tendency of the Bio-Tech students to disagree with the arguable claim does not increase following their participation in the Bio-Tech program compared to Control students, the former were better able to use arguments, and the number of arguments that focused on the experiment described in the article increased among the Bio-Tech students. The qualitative analysis supports the observed change in the type of arguments used by the Bio-Tech students before and after the intervention.

4. Discussion

Experiencing inquiry and gaining an appreciation of authentic scientific practices are key elements of science learning and teaching (National Research Council [NRC], 2012). The ability to critique is crucial in students' development of skills, abilities and understanding of scientific discourse and habits of mind (Berland & Reiser, 2009; Ford, 2008). In the study described herein, we explored possible development of students' ability to critique following their participation in the inquiry-oriented Bio-Tech program. No differences were observed in students' tendency to disagree with an arguable claim that was presented to them following the intervention between the Bio-Tech group and the Control group. Students from both groups appeared to be more in agreement with the arguable claim. This indicates that participation in the Bio-Tech program does not affect the students' ability to disagree more with an arguable claim. It may imply that developing students' ability to dispute and reject peer claims requires deeper and more explicit learning of critiquing. However, we found that participation in the Bio-Tech program leads to some improvement in students' ability to critique, mostly in their tendency to use more arguments and to critique experiments presented to them. Following participation in the program, the average number of arguments used in the pre- and post-questionnaires was sustained among the Bio-Tech group, in comparison to the Control group in which a significant decrease in the number of arguments used was observed in the post-questionnaires. This indicates that participation in the Bio-Tech program may have supported the students' argumentation and critiquing abilities.

The decrease in the average number of arguments used by the Control group might be explained by the fact that they were already familiar with the article presented in the questionnaire and they refrained from seriously engaging in answering the questionnaire. This may indicate that the ability and dedication of the Bio-Tech students to engage in critique about a topic that was already introduced in earlier experience have improved.

Furthermore, students of the Bio-Tech program tended to focus more on the experiment that was described in the article in their answers. This indicates that the Bio-Tech students improved some of their ability to critique and implies the possible development of this ability following participation in the Bio-Tech program.

Our results partially correlate with those presented by Ford (2012), who showed that students who focus on learning to critique while practicing an inquiry-oriented scientific activity improve their peer-review practice and their reasoning and argumentation abilities. The Bio-Tech students demonstrated development of their ability to critique, mostly enhancing the number of arguments used and the use of arguments related to the experimental process and method compared to the Control group. It should be noted that the Bio-Tech students' tendency to disagree with an arguable claim did not increase compared to students from the Control group, unlike the students who participated in Ford's Research (Ford, 2012).

Further research and analysis is required for a full understanding and appreciation of the development of students' ability to critique in the course of participation in inquiry-oriented programs. Deeper examination of the development of the ability to critique by inquiry-oriented students is required, due the relatively small number of students who participated in this research and the limited number of differences between the groups that were found. Our aim is to further analyze the development of students' ability to critique, to explore the students' long-term learning of critiquing and other abilities of the authentic scientific practice and to examine the learning of these abilities in other inquiry-oriented programs. We also plan to further and more deeply explore the development of students' ability to critique while participating in the Bio-Tech program, focusing on their ability to critique their own and their peers' research processes.

REFERENCES

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765-793.

Berland, L. K., & McNeill, K. L. (2009). Making sense of argumentation and explanation. *Science Education*, *93*(1), 26-55.

Berland, L. K., & Reiser, B. J. (2011). Classroom communities' adaptations of the practice of scientific argumentation. *Science Education*, *95*(2), 191-216.

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

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. B. R. J. Sternberg (Ed.), *Teaching thinking skills: Theory and practice* (pp. 9-26). New York, NY, US: W H Freeman/Times Books/Henry Holt & Co.

European Commission. (2007). Science Education Now: A renewed pedagogy for the future of Europe. Brussels, Belgium: Directorate-General for research: Science, economy and society.

Ford, M. J. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404-423.

Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction*, *30*(3), 207-245.

Israeli Ministry of Education. (2005). Syllabus of biotechnological studies (10^{th}-12^{th} Grade). Jerusalem, Israel: State of Israel Ministry of Education Curriculum Center.

National Research Council [NRC]. (2000). *Inquiry and the national science education standards*. Washington, DC: National Academic Press.

National Research Council [NRC]. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academic Press.

National Research Council [NRC]. (2012). A framework for K-12 science education: Practices, crosscutting concepts, and core ideas. Washington, DC: National Academic Press.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, *328* (5977), 463-466.

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

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

Siegel, S., & Castellan, N. J. J. (1988). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Stolarsky Ben-Nun, M., & Yarden, A. (2009). Learning molecular genetics in teacher-led outreach laboratories. *Journal of Biological Education*, *44*(1), 19-25.

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, *1*(6), 80-83.

Yarden, A. Brill, G., & Falk, H. (2001). Primary literature as a basis for a high-school biology curriculum. *Journal of Biological Education*, *35*(4) 190-195.