# INVESTIGATING THE EFFECTIVENESS OF ARGUMENT-BASED INQUIRY ON 6TH GRADE STUDENTS' SCIENTIFIC LITERACY AND PORTRAYING THEIR ARGUMENTATION SCHEMES AND ENGAGEMENT IN ARGUMENTATION PROCESS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MEHMET ŞEN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR PHILOSOPHY
IN
THE DEPARTMENT OF ELEMENTARY EDUCATION

MARCH 2021

Approval of the thesis:

**INVESTIGATING THE EFFECTIVENESS OF ARGUMENT-BASED INQUIRY ON 6TH GRADE STUDENTS' SCIENTIFIC LITERACY AND PORTRAYING THEIR ARGUMENTATION SCHEMES AND ENGAGEMENT IN ARGUMENTATION PROCESS**

submitted by **MEHMET ŞEN** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Elementary Education, the Graduate School of Social Sciences of Middle East Technical University** by,

Prof. Dr. Yaşar KONDAKÇI
Dean
Graduate School of Social Sciences ____________

Assoc. Prof. Dr. Elvan ŞAHİN
Head of Department
Department of Elementary Education ____________

Prof. Dr. Semra SUNGUR
Supervisor
Department of Math and Science Education ____________

Prof. Dr. Ceren ÖZTEKİN
Co-Supervisor
Department of Math and Science Education ____________

**Examining Committee Members:**

Prof. Dr. Esen UZUNTİRYAKİ KONDAKÇI (Head of the Examining Committee)
Middle East Technical University
Department of Math and Science Education ____________

Prof. Dr. Semra SUNGUR (Supervisor)
Middle East Technical University
Department of Math and Science Education ____________

Prof. Dr. Jale ÇAKIROĞLU
Middle East Technical University
Department of Math and Science Education ____________

Assoc. Prof. Dr. Yasemin ÖZDEM YILMAZ
Muğla Sıtkı Koçman University
Department of Primary Education ____________

Assoc. Prof. Dr. Sevgi KINGIR
Hacettepe University
Department of Primary Education ____________

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

**Name, Last Name:** Mehmet ŞEN

**Signature:**

## ABSTRACT

INVESTIGATING THE EFFECTIVENESS OF ARGUMENT-BASED INQUIRY ON 6TH GRADE STUDENTS' SCIENTIFIC LITERACY AND PORTRAYING THEIR ARGUMENTATION SCHEMES AND ENGAGEMENT IN ARGUMENTATION PROCESS


ŞEN, Mehmet
Ph.D., The Department of ELEMENTARY EDUCATION
Supervisor: Prof. Dr. Semra SUNGUR
Co-supervisor: Prof. Dr. Ceren ÖZTEKİN



March 2021, 424 pages



This study examined Argument Based Inquiry impact on 6th grade students' scientific literacy including content knowledge, epistemological beliefs, and science process skills and also portrayed students' argumentation schemes and engagement in argumentation process. Totally, 71 students participated in quantitative part and 35 students participated in qualitative part. Content knowledge tests, epistemological belief questionnaire and science process skill test were used as pre-test and post-test to collect quantitative data. Qualitative data was collected through classroom observations. Data were analyzed by within subject repeated MANOVA, and qualitative data were analyzed using constant comparative analysis. Among the scientific literacy components only participants' content knowledge increased significantly, but epistemological beliefs and science process skills did not change. Five assertions were constructed for students' argumentation schemes. First of all, number of argumentation schemes types increased over time. Secondly, students preferred some argumentation schemes more. Thirdly, data type affected use of argumentation schemes. Next, more active class used more argumentation schemes.

Lastly, most used argumentation schemes varied between classes. Similarly, eight assertions were proposed about students' engagement. Accordingly, students engaged argumentation at most when using evidence cards. They engaged less when they do similar activities. Ideal conditions inhibited their engagement. Students mostly used expositional comments. If students have prior knowledge and evidence cards, they used oppositional comments. Students used information seeking when conducting experiment. Lastly, students did not use co-construction frequently. Finally, more active class engaged argumentation when doing experiment, but less active class engaged argumentation more when using evidence cards. Findings are discussed in detail and suggestions are provided. Moreover, discussion part provides further information about scientific literacy, argumentation schemes and students' engagement in argumentation process.

**Keywords**: Argument Based Inquiry, Scientific Literacy, Argumentation Schemes, Engagement in Argumentation Process, Middle School

# ÖZ

ARGÜMAN TABANLI BİLİM ÖĞRENME'NİN 6. SINIF ÖĞRENCİLERİNİN BİLİMSEL OKURYAZARLIĞI ÜZERİNE ETKİSİNİN İNCELENMESİ VE ÖĞRENCİLERİN ARGÜMANTASYON ŞEMALARININ VE ARGÜMANTASYON SÜRECİNE KATILIMININ ORTAYA KONMASI

ŞEN, Mehmet
Doktora, İlköğretim Bölümü
Tez Yöneticisi: Prof. Dr. Semra SUNGUR
Ortak Tez Yöneticisi: Prof. Dr. Ceren ÖZTEKİN

Mart 2021, 424 sayfa

Bu çalışma Argüman tabanlı Bilim Öğrenmenin 6. Sınıf öğrencilerinin fen okuryazarlığı boyutlarına etkisini incelemekte ve öğrencilerin argümantasyon şemaları ile argümantasyon sürecine katılımlarını ortaya koymaktadır. Çalışmanın nicel kısmına 71, nitel kısmına 35 öğrenci katılmıştır. Nicel veri için içerik alan bilgisi testleri, epistemolojik inanç ölçeği ve bilimsel süreç becerileri testi ön-test ve son-test olarak kullanılmıştır. Nitel veriler sınıf gözlemleri ile toplanmıştır. Nicel veri analizinde tekrarlayan MANOVA, nitel veri analizinde sürekli karşılaştırmalı analiz kullanılmıştır. Fen okuryazarlığı boyutları göz önüne alındığında sadece katılımcıların içerik bilgisi anlamlı ölçüde artarken, epistemolojik inançları ve bilimsel süreç becerileri değişmemiştir. Argümantasyon şemaları ile ilgili beş iddia ortaya atılmıştır. Buna göre; kullanılan argümantasyon şeması çeşidi sayısı zamanla artmaktadır. Öğrenciler bazı argümantasyon şemalarını daha fazla kullanmıştır. Veri çeşidi argümantasyon şeması kullanımını etkilemektedir. Daha aktif sınıf daha fazla argümantasyon şeması kullanmıştır. En çok kullanılan argümantasyon şeması sınıftan sınıfa farklılık göstermiştir. Öğrencilerin argümantasyon sürecine katılımı ile ilgili olarak sekiz iddia ortaya atılmıştır. Buna göre; öğrenciler kanıt kağıdı kullanınca argümantasyona daha fazla katılmıştır. Benzer etkinlikler öğrencilerin argümantasyona katılımını azaltmıştır. Koşullar ideallere yaklaştıkça argümantasyona katılım azalmıştır. Öğrenciler argümantasyon sürecinde en çok kendi düşüncelerini ortaya koyma boyutuna yoğunlaşmıştır. Öğrenciler hem kanıt kağıdı kullanıp hemde konu hakkında ön bilgiye sahip olurlarsa başkalarının fikirlerini eleştirmeye çalışmıştır. Öğrenciler deney yaptıkları ünite başı haftalarında daha fazla soru sorma eğilimindedir. Öğrenciler çalışma boyunca beraber bilgi üretme sürecine çok dahil olmamışlardır. Son olarak, daha aktif sınıf deney yapılan haftalarda tartışırken az aktif sınıf kanıt kağıdı verilen haftalarda tartışmıştır. Sonuçlar detaylı olarak tartışılmış ve ilgili öneriler sunulmuştur. Ayrıca tartışma bölümünde fen okuryazarlığı, argümantasyon şemaları ve öğrencilerin argümantasyon sürecine katılımı ile ilgili detaylı bilgi sunulmaktadır.

**Anahtar Kelimeler**: Argüman Tabanlı Bilim Öğrenme, Fen Okuryazarlığı, Argümantasyon Şeması, Argümantasyon Sürecine katılım, Ortaokul

*To*

My Family

## ACKNOWLEDGMENTS

First of all, I would like to thank to my supervisor Prof. Dr. Semra Sungur and my co-supervisor Prof. Dr. Ceren Öztekin for their encouragement, guidance, suggestions and feedbacks in these six years. I have learnt how a scientific research is planned, conducted and reported with your guidance. Secondly, I would like to thank to my committee members who are Prof. Dr. Esen Uzuntiryaki-Kondakçı, Prof. Dr. Jale Çakıroğlu, Assoc. Prof. Dr. Yasemin Özdem Yılmaz and Assoc. Prof. Dr. Sevgi Kıngır for their valuable feedbacks and comments for my thesis. Next, I would like to thank to my other teachers Prof. Dr. Özgül Yılmaz Tüzün, Prof. Dr. Gaye Teksöz, and Assoc. Prof. Dr. Elvan Şahin for their motivation in my thesis process.

I also want to thank my research team who assisted me in this long process. Therefore, I would like to thank Dinçer Çam, Deniz Turan, Cansu Başak Uygun, implementing teacher M. K., and pre-service teachers who assisted implementation of the treatment in pilot and main study. Thanks to your cooperation, we worked together more than two years. I also want to thank to more than 100 hundred $6^{th}$ grade students who participated my research. In analysis process, I discussed with Deniz and we analyzed the data together. Deniz, you added quality to my analysis. I also want to thank you for this reason.

I would like to thank to my family members; Necati Şen, İfakat Şen, Semiha Şen, Semih Koray Şen, Semiha Şen-Kaya, Egemen Kaya, and Erhan Kaya for their support in this process.

During this process, my office mates always supported me. Therefore, I would like to thank to Sinem Demirci, her husband Levent, and their daughter Yağmur, Seçil Cengizoğlu, her husband Güçlü Cengizoğlu, and their daughter İpek, Gözde Kaplan Can, her husband Cem and their daughter Ada. I wish you best in your life with your family.

I also would like to thank to my buddy Turan Gülçiçek and his wife Merve. I wish you best in your life and thank you very much for your support.

Next, I would like to thank to my close friends Selim Özdem, Ali Engin, Nurgül Turgul, Defne Güliz Öztürk, and Ferhat Ünaldı. You always supported and motivated me in this process. When I disappointed, your advices were very helpful. I am very lucky guy to have friends like you.

I also would like to thank to my elder friends and teachers, so I want to thank to Dekant Kıran, Betül Ekiz Kıran, Işıl İşler-Baykal, Erdinç İşbilir, Ramazan Sak, and İkbal Tuba Sak. Your experiences and feedbacks always opened my sight. Thank you very much for your help.

I also would like to thank to research assistants in our faculty who are Gamze Çetinkaya Aydın, Ceren Başer, Coşkun Aykut, Mine Tanrısevdi, Olcay Ketenci, Ezgi Şenyurt, Sabiha Üzüm, Merve Arslan, Tuğçe Esra Terzioğlu, Havva Ayça Alan, and Celal İler.

**TABLE OF CONTENTS**

APPENDICES

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ABI | Argument Based Inquiry |
| MHCKT | Matter and Heat Content Knowledge Test |
| ECKT | Electricity Content Knowledge Test |
| EBQ | Epistemological Beliefs Questionnaire |
| SPS | Science Process Skills |
| MONE | Ministry of National Education |
| NGSS | Next Generation Science Standards |

# CHAPTER 1

# INTRODUCTION

## 1.1. Scientific Literacy

General aim of science education is to make students scientifically literate (Roberts, 2007). However, the philosophy underlying scientific literacy is complex and vague. According to Roberts (2007), the term "Scientific Literacy" gained importance in second half of the 20th century. The general aim was to make students potential scientists for the future. Therefore, important characteristics of scientists were looked for to determine whether someone is scientifically literate or not. In other word, who acted like scientists in their life were thought as scientifically literate person according to these initial attempts on scientific literacy. This view was named as Vision-1 (Roberts, 2007).

In line with this, some variables which are at the centre of science can be indicator for scientific literacy. For example, Pella, O'Hearn and Gale (1966) reported that scientific literate person should know basic concepts in science. Therefore, content knowledge can be indicator for scientific literacy. Content knowledge is the product which is the result of scientific enterprise. In this study, content knowledge is defined as academic content of a discipline (e.g., physics, biology, chemistry). This knowledge includes how to write a scientific explanation, understanding of a domain in a discipline, relationship between domains, and relationship between topics and concepts in that domain (Carlson & Daehler, 2019).

Content knowledge is the product which is the result of scientific enterprise (Bunterm et al., 2014). However, vision-1 not only deals with product of science, it also considers process of science (Robert, 2007). In line with this, scientific literate person engages in scientific practice through scientific talking and writing. By this way, students learn better how scientific knowledge is produced and what characteristics that scientific knowledge has (Jimenez-Aleixandre, & Erduran, 2008). In other words, students' epistemological beliefs can be another indicator for scientific literacy. Epistemological beliefs in this study are defined as the belief system including source, certainty, development, and justification. According to Conley, Pintrich, Vekiri, and Harrison (2004) sources of knowledge deal with where scientific knowledge comes from. Source can be people who conduct scientific research or outside of the people. Certainty is about whether scientific knowledge is certain or uncertain. Development is about whether scientific knowledge changes or not and justification emphasizes the role of experiment, testing and evidence in science. The beliefs consistent with constructivist ideas such as scientific knowledge can change based on evidence is hierarchically better than the positivist beliefs like scientific knowledge is certain and does not change (Conley et al., 2004).

As epistemological beliefs represent students' belief systems about scientific knowledge and the process that this knowledge is produced; science process skills are also at the centre of process of science and it is related with scientific literacy (i.e. vision-1) because students are expected to use science process skills while they engage in scientific enterprise before producing knowledge (Harlen, 1999). Science process skills can be defined as science related skills (Padilla, 1990). More specifically, science process skills are the results of the classification of the science's intellectual tools (Sanderson & Kratochvil, 1971). Accordingly, science process skills can be either basic process skills or integrated process skills. Basic process skills are observing, classifying, using numbers, measuring, using space/time relationships, communicating, predicting and inferring and these skills are taught in primary level. On the other hand, integrated process skills are defining operationally, formulating hypotheses, interpreting data, controlling variables, and experimenting and these skills are taught in middle school (Sanderson & Kratochvil, 1971). Participants of the current study is 6$^{th}$ grade students (i.e. middle school level), therefore, integrated process skills are decided to be the main focus of the study.

In sum, in this study, vision-1 for scientific literacy was adopted because vision-1 emphasizes product and process of science. Therefore, content of this study regarding scientific literacy includes content knowledge, epistemological beliefs and science process skills consistent with vision-1 for scientific literacy.

### 1.2. Argumentation

Although general aim of science education is to make students scientifically literate, there is no specific suggestion about which approach should be used to make students scientifically literate.  On the other hand, Lederman and Lederman (2012) claimed that students learn science best through inquiry oriented approach because they actively do science in this approach.

When people make inquiry, the following processes are observed: they ask a question at the beginning of the scientific investigation. There is no single scientific method to be followed because question asked at the beginning determines the method in scientific investigation. In other word, question guides the inquiry process. Although scientists follow the same process, they can reach different results. Likewise, inquiry process can affect the results. In this process, research conclusion has to be consistent with data. All data is not counted as evidence in this process and this make these two entities different from each other. Finally, scientists reach explanations for the question asked at the beginning. These explanations are developed from the sum of collected data and scientists' background knowledge (Lederman & Lederman, 2012).

Although inquiry is necessary, it is not sufficient to reach goals about scientific literacy. For example, Conley et al. (2004) examined the effect of inquiry approach on 5$^{th}$ grade students' epistemological beliefs, but this approach did not improve some components of epistemological beliefs. Conley et al. (2004) explained this result based on lack of argumentation. Accordingly, if inquiry is accompanied with argumentation, components of epistemological beliefs might improve more.

In this point, argumentation in science can be described. Argumentation in science is a process where competition, collaboration and negotiation occurs (Cavagnetto, 2010). While argumentation is a process in which knowledge is constructed, its ending product is argument (Jimenez-Aleixandre & Erduran, 2008). When students engage in argumentation process and produce arguments, they use some entities which are claim, data, warrant, qualifier, backing, and rebuttal. Accordingly, Claim is the assertion. Data is the fact that supports claim. Warrant is the fact that connects data and claim. Qualifier shows to what extent claim is suitable. Backings strengthen the warrant and increase trustworthiness. Rebuttals are the exceptional situations where claim is not true (Toulmin, 1958).

Previous argumentation studies were mainly conducted to reveal participants' argument quality (i.e., product of argumentation process) and there was less emphasis on argumentation process (Sampson, Enderle, & Walker, 2012). Therefore, researchers know less about argumentation process compared with the quality of students arguments. Likewise, how students learn arguing, support and refine their argument is not clear (Kim & Song, 2006). Similarly, Sampson and Clark (2011) claimed that it is not well known in which condition students prefer to use co-construction of knowledge or oppositional comment. For example; Sampson and Clark (2011) found that students mostly prefer to use oppositional comments in argumentation process when they selected the best idea among alternative ones. On the other hand, Sampson and Clark (2011) added that students might prefer to use co-construction of knowledge in argumentation, if task was developing students' original idea. In conclusion, there is a need to study with students demonstrating different levels of engagement in argumentation process.

Furthermore, studies conducted to reveal argument quality was also problematic in earlier research (Duschl, 2007). Accordingly, this early studies focused on number of justification in an argument, quality of justification (Zohar & Nemet, 2002), number of rebuttal, quality of rebuttal (Erduran, Simon, & Osbourne, 2004) to decide quality of argument. However, whether participants' arguments reflect epistemic criteria in these studies were not clear (Sampson & Clark, 2006). According to Sampson and Clark (2006) an argument having high quality should met five epistemic criteria which are (a) examining the nature and quality of the knowledge claim, (b) examining how the claim is justified, (c) examining if a claim is accounts for all available evidence, (d) examining how the argument attempts to discount alternatives and (e) examining how epistemological references are used to coordinate claims and evidence. In this point, Duschl (2007) claimed that Walton's (1996) argumentation schemes can be useful to reveal argument quality because argumentation schemes are consistent with five epistemic criteria proposed by Sampson and Clark (2006). Argumentation schemes are forms of arguments and they show structures of common arguments used in daily life (Walton, Reed, & Macagno, 2008). There are 25 argumentation schemes (e.g., Argument from sign) that Walton (1996) explained. When people use argumentation schemes, they use presumptive reasoning which means that there is a lack of evidence. Therefore, these arguments are not strong. Although they have lack of evidence, they can still have some evidence and so they have weight to change the direction of argumentation process. Using these schemes, some tentative conclusions are reached. This is consistent with science and daily life because always there is lack of evidence and so certain conclusion cannot be reached. However, argumentation schemes are fallible and tentative because when new evidence emerge, argumentation schemes can be falsified and the tentative conclusion which were constructed on argumentation schemes can also be rejected.

### 1.3. Theoretical consistency between Argumentation and Scientific Literacy

Theoretically, argumentation in science classes supports scientific literacy components which are content knowledge, epistemological beliefs and science process skills. First of all, when students engage in argumentation, they open their ideas to the class. Everyone becomes aware of others' thinking. The knowledge is shared among students, and they can construct the knowledge through argumentation. In this process, students make apprenticeship to each other and they learn from others. Students can also understand their wrong ideas and their peers' wrong ideas and they can replace wrong information with scientific ones through examination of ideas in argumentation. By this way, students' content knowledge increases (Jimenez-Aleixandre & Erduran, 2008).

Second, students study like scientists in argumentation. When students engage in argumentation process, they can understand the construction of scientific knowledge and characteristics of that knowledge by becoming member of scientific community.

This means that students develop their epistemological beliefs through practice (Sandoval & Millwood, 2008). For example, Khishfe (2014) claimed that people use evidence to persuade others in argumentation and by this way, students can understand that evidence is important in science. Likewise, students understand that tentative conclusions are not certain in science because counterarguments can change tentative conclusions and direction of the argumentation depending on the evidence they carry (Khishfe, 2014).

Third, argumentation is also consistent with science process skills. Due to fact that students act like scientists in argumentation (Jimenez-Aleixandre, 2008), they are expected to do similar activities that scientists do in their work (e.g., collecting data). By engaging science practice in argumentation, it is expectable that students improve their science process skills. For example, students formulate hypothesis prior to experiment. They identify the variables in their experimental design. Through discussing with others, they can produce some operational definitions to the variables. Likewise, they can discuss on the results of their experiments and they can interpret the result. The more they engage in scientific practice in argumentation, the more they use and improve their thinking skills related with science (i.e. science process skills).

### 1.4. Characteristic of an Argumentation Class

After presenting the theoretical consistency between argumentation and scientific literacy, it is meaningful to discuss the characteristics of a class where argumentation occurs. First of all, argumentation aligns with social constructivism because students construct the knowledge in both cases. Therefore, argumentation class is similar to other classes where constructivism is adopted (Jimenez-Aleixandre, 2008). Characteristics of argumentation class include student, teacher, curriculum, assessment, metacognition and communication components (Jimenez-Aleixandre, 2008).

In argumentation based lessons, students should be responsible for their own learning constructing the knowledge interacting with peers. They are expected to justify their claims and propose solutions to problems. They should consider alternative ideas and analyze them. When they propose ideas, they must use evidence. They should also be capable of distinguishing weak arguments from strong ones. During the argumentation practice, students should discuss solutions, make hypothesis, test them, and report the results (Jimenez-Aleixandre, 2008). Thus, in an argumentation class, students become metacognitively active:  when students propose their ideas, they become aware of their ideas' strengths and weaknesses (i.e., monitoring). Then, they are expected to adjust their ideas in order to eliminate their ideas' weaknesses (i.e., regulation). Moreover, students can explain the reasons about why their ideas changed over time (Jimenez-Aleixandre, 2008). In addition, interaction and dialogues are central in an argumentation class. All students are invited to this communication. Knowledge is produced as results of the interactions among students. During this process, students discuss, ask questions, evaluate their ideas and criticize others' ideas. In this interactive context, students try to persuade each other and there should be consensus at the end (Jimenez-Aleixandre, 2008).

As students' role, teacher role is also clear in argumentation lessons. According to Jimenez-Aleixandre (2008) teacher is not the authority in class; however, if students reach a false knowledge and they are in consensus about this issue, teacher contributes to discussion and points out new perspectives to be discussed. In this point, teacher acts as a facilitator. While teacher guide discussion, s/he also considers the curricular objectives. When students' arguments diverge from objectives, teacher may ask students to connect their arguments to the topic under investigation. Throughout the argumentation process, teacher encourages students to use evidence in claims and ask questions to others. Teachers should also tell students the characteristics of good arguments (e.g. use of justifications, rebuttals) and argumentation based lessons (e.g., talking, listening, debating). Finally, teacher should encourage students to change their ideas when their ideas do not match with evidence or scientific theories.

Curriculum is as important as teachers and students in an argumentation based class. It is difficult to conduct argumentation lessons in a curriculum where inquiry is not embedded (Jimenez-Aleixandre, 2008). In inquiry based curriculum, it is expected that students should act as scientists. Students should engage in authentic activities which is difficult to be implemented in regular classes. Curriculum should focus on the real life problems that students make research on it. Curricular activities should allow students to reach different conclusions, by this way students discuss these conclusions more. There should be only few numbers of goals or objectives in argumentation lessons. By this way, students focus on claims in selected topic and their evaluation process.

Argumentation lesson is not limited with student, teacher and the curriculum: assessment is also inevitable Teacher must also get feedback from the students regarding the efficiency of the lesson. In addition, teacher should prepare criteria for evaluation of students' engagement in argumentation and their product (e.g., quality of argument, content knowledge) and these criteria should be shared with students (Jimenez-Aleixandre, 2008). Throughout the process, students might prepare portfolios consisting students' old and new claims. Students make reflection by comparing these old and new claims, so they understand how their ideas changed and developed. Types of assessments should also vary because students can express themselves better when there is more than one type of assessment.

Characteristics of argumentation class were reflected in this study. These characteristics contributed on validity of this research also. For example; students were responsible for their learning and they were expected to interact with their peers as part of the student characteristics of argumentation class. Teacher role was adopted as facilitator in this study. Therefore, the evidence and robust arguments were more important than what teacher said in argumentation lessons. Curriculum characteristics of argumentation class also shaped current study. Accordingly, argumentation classes should include few but comprehensive objectives to sustain argumentation classes. Therefore, comprehensive and few numbers of objectives were selected in this argumentation study. Alternative types of assessment were also selected to demonstrate characteristics of argumentation class. For example; students wrote their initial and final arguments on their work sheets throughout the process. By this way, they were able to know their development. Likewise, they wrote what they learn in each lesson and they made reflections. Similarly, peer questions and teacher questions fed students’ understandings and revealed students’ wrong ideas. After listening these questions, students had chance to correct their wrong knowledge.

### 1.5. Researcher Philosophy

Previous research on argumentation mainly adopted social constructivism (Jimenez-Aleixandre, 2008; Kind, Kind, Hofstein, & Wilson, 2011; Perry & Dockett, 1998; Sampson & Clark, 2011; Walker & Sampson, 2013). However, philosophy adopted in this study is interactive constructivism, not social constructivism.

Interactive constructivism is similar with social-constructivism, but there are some differences. Knowledge is constructed in social plane according to social constructivism. Cultural values, beliefs and consensus affect the evaluation of knowledge in social constructivism; however, natural laws (e.g. physical laws like gravity) are not as important as cultural values, beliefs and consensus in social constructivism. On the other hand, nature has an active role in deciding whether something is true or not according to interactive constructivism. While knowledge is constructed in social plane and people construct the knowledge through interaction, people have to consider also the consistency between their talking and nature rules in interactive constructivism. Thanks to interactive constructivism, personal views are connected to scientific views because nature becomes an authority to decide what is true. Interactive constructivism also suggests that science has two aspects which are social and material aspects. While material aspect includes questions for exploration, designing appropriate ways to answer question and conducting investigation with accuracy, social aspect of science includes public debate, interpretation of evidence, and knowledge claims (Cavagnetto, Hand, & Norton-Meier, 2010).

If social constructivism had been adopted in this study, students’ discussions could have been shaped by their experience, prior knowledge and evidence obtained from their experiments. However, this philosophy ignores natural laws. For example, students conducted their experiments in first two weeks about heat conductivity of different substances like metal cup and plastic cup. If social constructivism had been adopted in this study, students could persuade each other by saying plastic cup is better heat conductor than metal cup because their experiment results supported this data. However, this assertion is problematic for interactive constructivism because natural laws should also be considered when deciding whether something is true or false. Accordingly, after students constructed their arguments about comparison of different materials about their heat conductivity, students were asked to consider whether their results are consistent with nature in line with interactive constructivism in this study. After that, students told a natural law that is metals conduct heat better than non-metals and they integrated this natural law to their ongoing argumentation. By this way, it is thought that argumentation studies prepared based on interactive constructivism provides more valid results comparing with the argumentation studies depending on social constructivism.

### 1.6. Researcher Orientation towards Argumentation

Cavagnetto (2010) examined different aspects of argumentation interventions by reviewing previous argumentation studies. Researcher focused on nature of activity, emphasis of activity, and science aspect in activity in this review. This analysis resulted with three different orientations that researchers have towards argumentation. These orientations are learning of argument through immersion, teaching the structure of argument and emphasizing the interaction between science and society (Cavagnetto, 2010).

First orientation is immersion in science for learning scientific argument (i.e., Immersion). In immersion orientation, argumentation is not done at the end of inquiry process; it exists in all parts of inquiry. For example; students engage in argumentation while asking questions, conducting experiments, interpreting data, and constructing claims. The aim is to understand science principles and science practice. Researchers provide tools to students as scaffolding which assist students to adapt argumentation practice (Cavagnetto, 2010). Second orientation is teaching the structure of arguments (i.e. structure). In structure orientation, researchers explicitly teach the structures of argument (e.g., claim, data) at first. Then, they expect students to apply these structures in new contexts. Structure orientation generally asks students to explain the natural phenomena without conducting experiment. When students do not conduct experiment, argumentation is seen as the product of inquiry process and it is not part of inquiry (Cavagnetto, 2010). Third orientation involves the emphasis on the interaction between science and society (i.e., socio-scientific). In this orientation, there is a socio-scientific issue and students use their existent knowledge to solve issue in general. Moral, ethical, and political considerations are emphasized in socio-scientific orientation. Connecting values to science is more important than understanding scientific principles in this orientation. Argumentation is used as a tool in socio-scientific orientation to show how socio-cultural factors shape science (Cavagnetto, 2010).

Accordingly, immersion orientation was adopted in this argumentation study because of three main reasons. Firstly, vision-1 approach for scientific literacy was adopted as scientific literacy. According to vision-1, students act like scientists and they actively engage in process and product of science in this approach (Roberts, 2007). Similar to vision-1 scientific literacy, immersion orientation for argumentation focuses on doing science (e.g. experimenting) and constructing knowledge. Secondly, immersion orientation includes all elements of science (e.g. controlling variables). However, structure and socioscientific orientations do not provide opportunity for students to understand all elements of science (Cavagnetto, 2010). For example, students do not conduct experiments in these orientations. Therefore, immersion orientation is more advantageous than other two orientations. Thirdly, interactive constructivism was adopted as philosophy in this study. According to interactive constructivism, science has two aspects namely social and material aspect. While material aspect deals with question for exploration and conducting investigation, social aspect of science includes public debate. Both aspects of science mentioned in interactive constructivism were emphasized in immersion orientation. For example, students make investigation and discuss on this investigation in immersion orientation. On the other hand, other two orientations only deal with social aspect of science ignoring material aspect of science (Cavagnetto et al., 2010). Accordingly, role of nature in science is ignored when material aspect of science is ignored. Therefore, it is thought that immersion orientation provides more information about science. Because of these three reasons, immersion orientation was adopted in this study.

## 1.7 Theoretical Framework

In line with philosophy (i.e. interactive constructivism) and orientation towards argumentation (i.e. immersion) of the study, theoretical framework of the study is Argument Based Inquiry (ABI). As a specific example of ABI, SWH approach proposed by Hand and Keys (1999) was used in this study. SWH is not just a tool that engages students to argumentation process. SWH is examples of immersion orientation in which students conduct investigation to engage in argumentation with peers (Cavagnetto, 2010; Chen, Hand, & Park, 2016).

Because students have difficulties when they engage in immersive approaches (i.e., Argument-Based Inquiry), ABI approach provides students with Science Writing Heuristic (i.e. SWH) templates facilitating students' adaptation and participation in argumentation practice (Chen et al., 2016).

Accordingly, the SWH has two templates namely teacher template and student template. The teacher template focuses on pedagogical issues to assist teachers in implementation of the SWH. On the other hand, the student template is prepared for assisting students in lab investigations. SWH brings writing aspect of science/ scientists in class and it is a pedagogical tool for students to encourage scientific reasoning in class (Hand, Norton-Meier, Gunel, & Akkus, 2016).

Teacher template outlines series of activities that promote student thinking. First of all, students prepare a concept map to show their prior understandings according to teacher template. Revealing students' prior knowledge facilitates asking questions before investigation. Secondly, individual pre-lab activities start. Students observe, explore, make brainstorm, and write a question. Thirdly, lab activities start. If lab activity is complex, teacher can inform students about procedure. If activity is not complex, students can design their own procedure. Lab activities prepare students to negotiations. Negotiations are step 4-5-6-7 where step 4 is individual negotiation, step 5 is peer negotiation, step 6 is textbook, teacher or authority negotiation and step 7 is negotiation as reflect, elaborate and share knowledge. Students repeat different steps throughout the inquiry process (Hand & Keys, 1999). Throughout the research, these phases were followed in this study. However, first phase which asks students to draw concept map was removed. Instead of concept mapping to understand students' prior knowledge, whole class discussions about daily life connections of the topics were added to reveal students' prior knowledge.

Second component of SWH is student template. Student template assists students to explain observed phenomenon. If students carefully complete the student template, they connect question, evidence, and claim with each other in their investigation. In student template, students write their question at first. Secondly, they write their testing procedure and they record their observations in third prompt. Prompt 2 and 3 are like traditional laboratory activities. In prompt 4-5, students write their claims and their reasons by constructing deeper understanding. In 6$^{th}$ prompt, results are compared with authorities. If results conflict with authority, students should negotiate with authority. Lastly, students write what they learn (Hand & Keys, 1999).

It should be noted that ABI approach is more than providing students SWH templates. By using ABI, students' initial ideas are elicited through pre-discussions and scientific investigations. Students get confused when they see their initial ideas are wrong, and they become motivated to correct their wrong ideas. Therefore, ABI aligns with conceptual change approach (Aydeniz, Pabuccu, Cetin, & Kaya, 2012). Likewise, students ask questions to each other and examine their ideas throughout the process (Aydeniz et al., 2012). Moreover, students formulate hypothesis, conduct experiments, engage in small group discussion and whole class discussions, and compare their findings with authorities (Hand & Keys, 1999). In conclusion, it can be said that SWH student template is only one aspect of ABI approach.

## 1.8. Purpose of the Study

This study has three main purposes in general. Firstly, this study aims to understand impact of Argument Based Inquiry (ABI) approach on middle school students' scientific literacy in terms of content knowledge, epistemological beliefs and science process skills. While this study focusing on the impact of ABI on scientific literacy, it also aims to exhibit how students engage in argumentation process during ABI treatment. Therefore, the study aims to portray students' engagement in argumentation process as second purpose. Likewise, this study does not only focus on process of argumentation, but it also deals with the product of argumentation which is argument. Hence, third purpose of the study is to reveal middle school students' argumentation schemes showing participants' arguments.

### 1.9. Significance of the Study

Previous ABI treatments mainly focused on examining the impact of ABI treatment on participants' content knowledge and learning (Cronje, Murray, Rohlinger, & Wellnitz, 2013; Greenbowe, Rudd, & Hand, 2007; Hand, Wallace & Yang, 2004; Hand et al., 2016; Hohenshell & Hand, 2006; Kingir, Geban & Gunel, 2013; Taylor, Tseng, Murillo, Therrien, & Hand, 2018). On the other hand, there is no study examining the impact of ABI treatment on middle school students' science process skills. Likewise, there is only one study examining the impact of ABI on middle school students' epistemological beliefs (Tucel, 2016). To sum up, it can be said that previous research did not examine the effect of ABI on scientific literacy which includes content knowledge, epistemological beliefs and science process skills empirically although there is a theoretical consistency between ABI and scientific literacy. For example, Jimenez-Aleixandre and Erduran (2008) claimed that students read different sources before reaching an argument, criticize sources, discuss with others. By this way, they learn scientific talking and writing improving their scientific literacy. However, the researcher of this dissertation did not come across with any empirical investigating the effect of ABI on scientific literacy.

On the other hand, this ABI study aims to improve these three components of scientific literacy. Actually, this study has some specific significance regarding the components of scientific literacy. First of all, this study aims to improve participants' content knowledge that is one of the aims of science curriculum [Ministry of National Education] (MONE, 2018). The content knowledge that students improve assists them to solve daily life problems (MONE, 2018). Similarly, new learnings are constructed on students' content knowledge. Therefore, improving students' content knowledge for this grade level ($6^{th}$ grade) can be helpful for further learning in following years (e.g. high school). Secondly, improving students' epistemological beliefs which are another component of scientific literacy is important. According to Next Generation Science Standards (NGSS, 2013), science is both a set of practices and accumulation of knowledge. Scientific literate students are expected to engage in science practice. However, engaging in scientific practice is not sufficient. Students should also know the characteristics of science practice. If these characteristics are not well known, contributions of scientific practice cannot be learnt. Characteristics of scientific practice represent epistemic knowledge that is the knowledge of constructs and values intrinsic to science. If students improve their epistemological beliefs, they can understand observation, hypothesis, inference, model, theory, claim and they can distinguish these characteristics of science (NGSS, 2013). By this way, students can understand both science practice and its product (content knowledge). Thirdly, the study aims to improve students' science process skills as scientific literacy components. Improving science process skills are crucial for students because these skills are used when scientists conduct their investigations to understand nature (MONE, 2018). Although science process skills are not explicitly referred in NGSS (2013) report, there is a consistency between science process skills and science practice. While science process skills focus on the skills that scientists use in their work, science practice focuses on performance including skills and knowledge. However, there is an overlap between science practice and process skills. For example, both themes focus on planning and carrying out investigations, endeavor and interpreting data, obtaining, evaluating and communicating information. Therefore, advantages of improving students' science practice can be also sign for benefits of improving science process skills. According to NGSS (2013), students can realize how science works and produce knowledge when they engage in science practice. Likewise, students appreciate science because students observe that scientific approaches investigate, model and explain natural phenomena. Similarly, engaging in scientific practice increases students' curiosity, interest, and motivation towards science. Furthermore, engaging in science practice assists students to understand crosscutting concepts and disciplinary ideas. Lastly, students can understand that World can change through human endeavor when scientists engage in science practice (NGSS, 2013). All these contributions of engaging in science practice are also true for science process skills and therefore, it is important to examine whether students' science process skills improve through ABI treatment.

Another significance of the study is about the use of argumentation schemes in this study. Sampson and Clark (2006) claimed that previous research conducting to reveal argument quality did not meet epistemic criteria (e.g. examination how the argument attempts to discount alternatives) and Duschl (2007) added that only Walton's (1996) argumentation schemes meet these epistemic criteria. Therefore, Duschl (2007) called researchers to analyze arguments using argumentation schemes. However, there are a few studies using argumentation schemes to analyze quality of arguments. For example, there is one ABI study examining students' argument quality using argumentation schemes, but this study was held with pre-service teachers (Ozdem, Ertepınar, Cakıroglu, & Erduran; 2013) and there is no ABI study examining quality of students' arguments by using argumentation schemes in middle school level. On the other hand, this ABI study aims to analyze middle school students' argument quality using argumentation schemes. Focusing on argumentation schemes in middle school level is important because use of argumentation schemes is indicator of students' reasoning. In other word, revealing students' argumentation schemes inform researchers about how students think. In ABI research, students are mentally active. By revealing argumentation schemes, their thinking process becomes clear and so researchers can understand better what is going on during argumentation.

Moreover, in his study, Duschl (2007) focused on argumentation schemes middle school students used and researcher found that middle school students used nine different argumentation schemes which are sign, position to know, evidence to hypothesis, cause to effect, analogy, commitment, expert opinion, corelation to cause, and consequences, but he did not use argument-based inquiry. On the other hand, this study used argument-based inquiry approach. In this approach, students may use much more types of argumentation schemes unlike Duschl (2007). For example, students can use argument from example when they reach a result in experiment and support this result with daily life example in ABI instruction. Likewise, students can use argument from gradualism in ABI instruction because students try to persuade others in whole class discussion part of ABI. One argument may not be sufficient to persuade others and students can connect different arguments to each other and they can construct one comprehensive argument reaching argument from gradualism. By this way, this comprehensive argument may be more convincing in whole class discussion. To sum up, ABI instruction promise use of various types of argumentation schemes. If this approach assists students to reveal various types of argumentation schemes, much more information about nature of middle school students' argumentation schemes can be learnt by this way.

Focusing on students' engagement in argumentation process is another significance of this study. Previous research on argumentation claimed that researchers mainly focused on quality of argument that is the product of argumentation process; however, there is little known about participants' engagement in argumentation process and there should be more studies focusing on students' engagement in argumentation process (Kim & Song, 2006; Sampson et al., 2012). Thus, this study can enhance our understandings about students' engagement in argumentation process in Turkish context. Studying on students' engagement in argumentation process can be meaningful because of some reasons. Firstly, argumentation is a process that students engage in. In this process, students compete, collaborate and negotiate with each other. If students' engagement in argumentation is investigated deeply, some new knowledge can reveal regarding how students compete, collaborate and negotiate with each other. Next, studying on students' engagement in argumentation process can show researchers in which conditions students compete with each other and in which conditions students collaborate. Likewise, the arguments that students use are the products of their engagement in argumentation process. Therefore, the possible trends or patterns obtained from students' engagement in argumentation process can be related with the arguments they produced. By this way, eliciting students' engagement in argumentation process can provide more information about the nature of students' arguments. For example, it is possible that students can criticize, defend and support arguments in their engagement to argumentation process. The argumentation schemes that students use in these different engagement types may differ. It is possible that some argumentation schemes are proposed more frequently when students criticize each other, and other argumentation schemes are used more frequently when students work in cooperation in their engagement in argumentation process. Therefore, increasing knowledge about students' engagement in argumentation process assists researcher to reach further understandings about the arguments students use. As connections between engagement in argumentation process and argumentation schemes that students use inform our understanding about argumentation, these two themes' connections with scientific literacy can also enhance our understanding further. For example, the study might show which argumentation schemes let students to reach scientific knowledge. By this way, the relation between argumentation schemes and content knowledge that is one aspect of scientific literacy may become obvious. Likewise, engagement components' link with content knowledge can increase our understanding about these two themes. For instance, students might reach scientific knowledge when they use one specific engagement component (e.g. oppositional comment) and students might not reach scientific knowledge when they use some other engagement components (e.g. co-construction of knowledge). If such instances are explored, teachers may monitor their students regarding the use of engagement component leading students to reach scientific knowledge. Similar links can be also available between argumentation schemes, engagement components, and other two aspects of scientific literacy which are epistemological beliefs and science process skills. For example, if students do not use argumentation schemes like argument from sign or argument from or argument from evidence to hypothesis, this might show that students have deficiency on their science process skills since students use their observation and inference skills when they use argument from sign (Duschl, 2007). Likewise, students use formulating hypothesis skill when they use argument from evidence to hypothesis schema.To sum up, argumentation schemes students used might inform their science process skills. In a similar vein, students' engagement in argumentation schemes might inform us about students' epistemological beliefs. For example, if students do not engage in argumentation and there is no frequent use of engagement components, this might show that students hold less sophisticated epistemological beliefs because students with sophisticated epistemological beliefs are expected to engage in argumentation process actively. Likewise, students' use of argumentation schemes can provide information about their epistemological beliefs. If students do not use scientific argumentation schemes like argument from sign, argument from evidence to hypothesis, and argument from corelation to cause; this might show that students do not have sophisticated epistemological beliefs and so they do not use scientific argumentation schemes. In other word, limited use of scientific argumentation schemes can be sign of less sophisticated epistemological beliefs. In conclusion, current study can both provide information about argumentation schemes, engagement in argumentation process, scientific literacy and these three themes possible connections. By this way, our understanding about these three important aspects of science education can increase further.

Current study also has a potential significance for researchers in the field of science education. In this study, it is hypothesized that argument based inquiry approach improves students' scientific literacy (vision-1) because of the theoretical link between ABI and scientific literacy components. If hypothesis is supported by empirical evidence (i.e. findings) in this study, researchers can conduct similar treatments with different samples having different characteristics for the generalizability of the current findings with an ultimate aim of enhancing students' scientific literacy. On the other hand, if this hypothesis is not supported in this study, researchers can focus on the reasons for why ABI treatment did not increase students' scientific literacy. By this way, researchers' understandings about ABI, scientific literacy and their relation increase.

Lastly, this study points out students' reactions, ideas, and performance in ABI treatment. For example, teachers can consider the most and the least used argumentation schemes in this study when they prepare an argument based inquiry lesson. Likewise, the components that students used the most and the least frequently in their engagement in argumentation process in this study can be informative for teachers and these components can shape teachers' preparation of ABI lessons.

Similarly, teachers can benefit from content of the each week's activities explained in detail in method chapter when they teach matter and heat and electricity units in $6^{th}$ grade level. Furthermore, difficulties that were encountered during implementation of ABI treatment in this study can be considered by teachers. If these difficulties are considered and eliminated in science classes, students benefit from ABI treatment more.

### 1.10. Research Questions

In line with significances of the study, this study has five main research questions:

1. Is there a change in $6^{th}$ grade students' content knowledge from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?

2. Is there a change in $6^{th}$ grade students' epistemological beliefs from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?

3. Is there a change in $6^{th}$ grade students' science process skills from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?

4. What is the nature of $6^{th}$ grade students' arguments when they are analyzed based on argumentation schemes?

5. What is the nature of $6^{th}$ grade students' engagement in argumentation process in ABI treatment?

### 1.11. Operational Definitions

*Scientific Literacy*: Scientific literacy is a general term that centralizes science as the main goal of science education according to vision-1. In this approach (vision-1), science related variables like content knowledge form scientific literacy. Scientifically literate people are expected to engage in scientific practice in this approach (Roberts, 2007).

*Content Knowledge*: Content knowledge is the product which is the result of scientific enterprise. In this study, content knowledge is defined as academic content of a discipline (e.g., physics, biology, chemistry). This knowledge includes how to write a scientific explanation, understanding of a domain in a discipline, relationship between domains, and relationship between topics and concepts in that domain (Carlson & Daehler, 2019). Students' conceptual understanding about matter and heat and electricity units taught in 6$^{th}$ grade levels were seen as their content knowledge in this study. Content knowledge was assessed by the Matter and Heat Content Knowledge Test and Electricity Content Knowledge Test which were prepared by researchers.

*Epistemological Beliefs*: Epistemological beliefs are the belief system including source, certainty, development, and justification as measured by the Epistemological Beliefs Questionnaire. According to Conley et al. (2004), sources of knowledge deals with where scientific knowledge comes from. Source can be people who conduct scientific research or outside of the people. Certainty is about whether scientific knowledge is certain or uncertain. Development is about whether scientific knowledge changes or not and justification emphasizes the role of experiment, testing and evidence in science.

*Science Process Skills*: Science process skills are the results of the classification of the science's intellectual tools. Accordingly, science process skills can be either basic process skills or integrated process skills. Basic process skills are observing, classifying, using numbers, measuring, using space/time relationships, communicating, predicting and inferring and these skills are taught in primary level. On the other hand, integrated process skills are defining operationally, formulating hypotheses, interpreting data, controlling variables, and experimenting and these skills are taught in middle school (Sanderson & Kratochvil, 1971). In this study, students' integrated process skills are considered as science process skills and measured by use of Test of Integrated Process Skills 2 (Burn, Okey, & Wise, 1985).

*Argumentation*: Argumentation in science is a process where competition, collaboration and negotiation occurs (Cavagnetto, 2010). Argumentation is done to understand nature by asserting supporting, defending, criticizing, adjusting claims in order to validate claims (Sampson et al., 2012).

*Argument*: While argumentation is a process in which knowledge is constructed, its ending product is argument (Jimenez-Aleixandre & Erduran, 2008).

*Argumentation Schemes*: Argumentation schemes are forms of arguments and they show structures of common arguments used in daily life (Walton et al., 2008). There are 25 argumentation schemes like argument from sign (Walton, 1996). When people use argumentation schemes, they use presumptive reasoning which means that there is a lack of evidence. Therefore, these arguments are not strong, but they are used to support main arguments. Argumentation schemes have weight and so they can change the direction of argumentation (Walton et al., 2008).

*Engagement in Argumentation Process*: Students' engagement in argumentation process represents nature and function of discussion when middle school students participate in argumentation in this study. Students' engagement in argumentation process was determined based on four codes obtained from Sampson and Clark's (2011) argumentation analysis. These codes are information seeking (e.g., requesting more information), expositional comment (e.g., proposing ideas), oppositional comments (e.g., challenge), and co-construction of knowledge (e.g., supporting ideas of others).

*Interactive Constructivism*: Interactive constructivism is a kind of constructivism, however, personal views are connected to scientific views because nature becomes an authority to decide what is true for interactive constructivism. Moreover, according to interactive constructivism, science has two aspects which are social and material aspects. While material aspect includes questions for exploration, designing appropriate ways to answer question and conducting investigation with accuracy, social aspect of science includes public debate, interpretation of evidence, and knowledge claims (Cavagnetto et al., 2010).

*Immersion Orientation*: In immersion orientation, argumentation is not done at the end of inquiry process; it exists in all parts of inquiry. For example; students engage in argumentation while they asking questions, conducting experiments, interpreting data, and constructing claims. The aim is to understand science principles and science practice (Cavagnetto, 2010).

*Argument Based Inquiry*: Argument Based inquiry is an immersion approach that facilitates students' engagement in argumentation. In this approach, students engage in pre-discussion, group discussion, whole class discussion, design and conduct their investigation, negotiate on their findings with others, compare their findings with authorities, and make reflection on their learning. While students engage in argumentation in ABI process, they also complete science writing heuristic (SWH) student template that facilitates students' engagement of argumentation and construction of arguments (Hand & Keys, 1999).

# CHAPTER 2

# LITERATURE

This chapter focuses on reviewing literature related with this study. At the beginning of the chapter, general information about scope of this research is shared. Next, the relation between argumentation and content knowledge are presented. Then, the relation between argumentation and epistemological beliefs are presented based on argumentation literature. After that, argumentation and science process skills' connections are explained. After the relationships between argumentation and scientific literacy components (e.g., content knowledge) are explained, studies about argumentation schemes are presented. At the end of this literature review, studies about students' engagement in argumentation process are reported.

## 2.1. Argumentation

Argumentation in science is a process where competition, collaboration and negotiation occurs (Cavagnetto, 2010). Argumentation is done to understand nature by asserting supporting, defending, criticizing, adjusting claims in order to validate claims (Sampson et al., 2012). Argumentation has three main characteristics. First, it has justification of knowledge claim. Second, it has persuasion. Third, there is a debate between groups (Jimenez-Aleixandre & Erduran, 2008). As a result of argumentation process or when participants engage in argumentation process, they produce their argument (Jimenez-Aleixandre & Erduran, 2008). On the other hand, some researchers use the words which are argumentation and argument interchangeably. For example, according to Walton (1996), an argument is conversation between two people to persuade each other. Considering related literature, in the current study, argumentation is conceptualized as a process and argument is as a product of this process.

Although it was not explicitly addressed, previous studies showed that argumentation is theoretically linked with scientific literacy (Jimenez-Aleixandre, 2008; Jimenez-Aleixandre & Erduran, 2008; Khishfe, 2014; Sandoval & Millwood, 2008). According to Roberts (2007), there are two different approaches regarding scientific literacy which are vision-1 and vision-2. In vision-1, scientific enterprise is focused. What scientists do when producing scientific knowledge and the product of this endeavour (content knowledge) is emphasized in vision-1. In other word, both process and product of science are important in vision-1. On the other hand, science and its relationship with society and science's contributions to solution of social problems are focused in vision-2 approach for scientific literacy. Vision-2 approach does not deal with what scientists do when producing scientific knowledge and vision-2 only deals with the product of science. Due to fact that vision-1 deals with both process of science and product of science, this study adopted vision-1 approach for scientific literacy.

As there are different approaches to scientific literacy, there are three different orientations that researchers adopted when they conduct argumentation studies which are immersion, socio-scientific and structural orientations (Cavagnetto, 2010). In immersion orientation, argumentation is a tool to understand science principles and practice. While students make investigation, this orientation embeds argumentation in intervention. Argumentation is not at the end of inquiry process, it exists in all parts of process. For example; students engage in argumentation while they are asking questions, conducting experiments, interpreting data, and constructing claims. By immersion, students learn science language over time. Two examples of the immersion approach are argument based inquiry (also known as science writing heuristics) (Hand et al., 2004) and argument driven inquiry (Walker, Sampson, & Zimmerman, 2011). On the other hand, socioscientific orientation aims to show how society factor affect science. Science and society are intertwined. Scientists debates; however, weight of scientists' debates are lighter than other components of society like ethics, religion, and culture (Cavagnetto, 2010). In 2004, Sadler analyzed studies adopting socioscientific orientation. This analysis showed that researchers focused on informal reasoning. Findings also showed that students' claims are unjustified.

They are unable to construct counterarguments, and cannot use scientific evidence. Students are not capable of analysing scientific arguments. Context is socio-scientific issue and students participate in this context and learn argumentation. Culminating activities like class debates and role play are so much in this orientation. Moral, ethical and political considerations are emphasized in this orientation. Constructing on understanding the scientific principles are not emphasized in this orientation. Connecting values to science content is important in this orientation. Next, Structure orientation focuses on whether structures of argument are adapted in different contexts. It includes communication and defending claims. Controlling variables, experiment trials, errors, data transformation, interpreting data do not exist in structure orientation. In structure orientation, argumentation is the product of inquiry process and it is not part of inquiry (Cavagnetto, 2010). When these three orientations are compared with each other, it is easily seen that only immersion approach focuses on both product and process of science. Moreover, immersion approach is consistent with vision-1 approach for scientific literacy. Therefore, immersion approach as argumentation orientation was adopted in this study. Due to fact that immersion orientation was adopted in this argumentation study, different approaches prepared in line with immersion approach are noticeable. These two approaches are argument based inquiry (Hand et al., 2004), and argument-driven inquiry (Walker et al., 2011). Argument driven inquiry was prepared in order to transform undergraduate level laboratories (Walker et al., 2011); however, focus of this study is examining middle school students' argumentation process and arguments. Therefore, argument driven inquiry was not selected as argumentation treatment in this study. On the other hand, science writing heuristic or argument based inquiry which has been applied in different education levels (see Chen et al., 2016, Pock, Burke, Greenbowe, & Hand, 2007) were selected as argumentation treatment.

In argument-based inquiry, students directly engage in inquiry process through asking questions, doing investigation, obtaining data, using evidence, proposing claims, and making discussion on their ideas (Chen et al., 2016). In other word, they engage in all activities that scientists do in scientific practice. Due to fact that students directly engage in scientific activities by doing argumentation, they do not need to learn components of argument. Therefore, components of the arguments are not cognitive load for students and students can focus on what they actually need to learn (e.g., content, skills). According to argument-based inquiry approach, there are two parts including teacher and student template. Teacher template outlines series of activities that promote student thinking. Second template is student template. Student template assists students to explain observed phenomenon (Hand & Keys, 1999). Details of teacher and student templates used in ABI approach were explained in previous chapter.

Both vision-1 approach for scientific literacy and immersion orientation for argumentation deals with both process of science and product of science. As vision-1 approach focuses on process of science and its product, this study focuses on students' content knowledge, epistemological beliefs and science process skills as scientific literacy components. In that point, content knowledge can be seen as product of science and epistemological beliefs and science process skills can be seen as the constructs used in process of science. Therefore, the studies considering the relations between argumentation and corresponding scientific literacy components are reviewed in the following first three parts. Similarly, immersion orientation for argumentation deals with both process and product of science. In this point, students' arguments' quality can be seen as product of scientific endeavour because they are proposed as a result of argumentation process and students' engagement in argumentation process can be seen as process of science. Therefore, following fourth part of literature review focuses on the studies about students' arguments' quality and last part focuses on the studies focusing on students' engagement in argumentation process.

Argumentation studies in science accelerated after 1980 (Cavagnetto, 2010; Chang, Chang, Tseng, 2010; Lee, Wu, & Tsai, 2009; Lin, Lin, Potvin,& Tsai, 2018; Tsai & Wen, 2005) and researchers mainly focused on quality of product (i.e. Argument) rather than argumentation process (Sampson et al., 2012). While determining the quality of an argument, researchers used argument components like justification and rebuttal (Erduran, 2007). For example, Zohar and Nemet (2002) removed warrant, backing, and qualifier from analysis and focused on justification while analysing quality of students' arguments. Justification is the reason that explains why data supported the claim. Whether content of the justification is scientifically correct or wrong determined the quality of arguments (Zohar & Nemet, 2002). On the other hand, Erduran et al., (2004) focused on rebuttal as sign of high quality argument because it informs people about the strengths of their arguments. However, Sampson and Clark (2008) reported that previous studies' results are problematic regarding argument quality because previous studies' way of analysis for argument quality did not meet the epistemic criteria. In their review, Sampson and Clark (2008) examined different approached to analyze participants' argument quality. In this review, Sampson and Clark (2008) claimed that same argument can be in high quality or low quality based on the framework that researchers use because different frameworks' structures, content and justification are different from each other. For example, some frameworks see justification as information, and they focus on whether there are some information exists in justification part of the argument. If the information exists, the argument is labelled as high quality. On the other hand, some frameworks (e.g., Lawson, 2003) thought justification as thinking process. The ones who adopt these frameworks focus on whether different aspects like evidence and prediction fits with each other to decide quality of arguments. Moreover, in an earlier review of analytical methods, Sampson and Clark (2006) claimed that analytical methods do not provide sufficient information about how participants' arguments reflect epistemic criteria. To make these epistemic criteria clearer, Sampson and Clark (2006) proposed five epistemic criteria which are examining the nature and quality of the knowledge claim, examining how the claim is justified, examining if a claim is accounts for all available evidence, examining how the argument attempts to discount alternatives and examining how epistemological references are used to coordinate claims and evidence. According to Duschl (2007), Walton's (1996) argumentation schemes fit with these five epistemic criteria. Therefore, it is thought that studies using argumentation schemes as analytical methods provide valid results about quality of arguments. Hence, studies including argumentation schemes are reviewed regarding argument quality results of previous studies.

Next, both argumentation process and product of argumentation (argument) are in scope of the study. According to Sampson et al. (2012), researchers focused on argumentation process less than its product. Therefore, it is possible that there is little known about argumentation process. In line with this, the studies examining participants' engagement in argumentation process to reveal what is known about students' argumentation process are presented after reviewing the studies focusing on argumentation schemes.

## 2.2. Argumentation and Content Knowledge

In this part, theoretical consistency between argumentation and content knowledge is emphasized at first. After that argumentation studies considering content knowledge are reviewed. Summary of the argumentation studies related with content knowledge are presented at the end of this part.

### 2.2.1. Theoretical Consistency between Argumentation and Content Knowledge.

Science education researchers support the view that argumentation foster students' content knowledge (e.g. Chin & Osborne, 2010; Dawson & Venville, 2013; Mendonça & Justi, 2014). In this dissertation, content knowledge is defined as academic content of a discipline (e.g., physics, biology, chemistry) (Carlson & Daehler, 2019). Since current dissertation focuses on the argument-based inquiry impact on selected variables (e.g., content knowledge), firstly theoretical consistency between argumentation and content knowledge is presented and secondly theoretical consistency between argument-based inquiry and content knowledge is presented.

According to Chin and Osborne (2010), people make reasoning on ideas when they engage in argumentation. People not only focus on why an idea is correct, but also examine why alternative ideas are wrong in argumentation. As a result, distinguishing the correct and wrong ideas considering the reasons facilitates conceptual understandings. Besides, theoretical consistency between argumentation and content knowledge was explained by Mendonça and Justi (2014) from different perspective. Accordingly, Mendonça and Justi (2014) explained the contribution of argumentation on content knowledge pointing out examination of evidence in models through argumentation. Mendonça and Justi (2014) claimed that evidence does not provide our understanding alone. Evidence is used to construct models. If evidence in a model works in reality, this makes model scientific. Whether an evidence works in reality is determined by justifications and refutations (e.g., examining evidence). By examining evidence through justifications and refutations, most suitable models explaining the scientific phenomenon are selected. These most suitable models represent the scientific knowledge. In conclusion, it can be said that by examining evidence in argumentation, people construct scientific models and they reach scientific knowledge. By this way, argumentation results with increasing conceptual understanding.

Similarly, Dawson and Venville (2013) pointed out evaluating data and discussing on it when they explained how argumentation fosters content knowledge. Accordingly, people understand the data better when evaluating it. Likewise, people discuss with each other when evaluating data. Both evaluating data and creating common understanding with others increase people content knowledge. As Dawson and Venville (2013) considered discussing on data as a mediator to increase content knowledge, Aydeniz et al., (2012) explained how argumentation supports content knowledge referring reasoning, conceptual change and metacognition. Accordingly, Aydeniz et al. (2012) thought that students elaborate their ideas, and ask questions to each other that requires rational explanations. This rationality leads better understandings and elimination of misconceptions. Regarding metacognition, for example, students organize their knowledge by use of written argumentation. This in turns increases students' metacognitive thinking. By this way, they become aware of their deficiencies. Explanations of peers in verbal argumentation and further thinking on their deficiencies assist students to increase their content knowledge. Aydeniz et al. (2012) also claimed that argumentation is similar to conceptual change approach because students restructure their initial ideas which assist students to consider their weaknesses and other variables that they did not pay attention before. Moreover, Aydeniz et al. (2012) pointed out the importance of peer questions that develop students' deeper understanding of the content.

As argumentation supports students' content knowledge, argument-based inquiry (i.e., ABI) is expected theoretically to support content knowledge too. In ABI process, students ask research questions, collect and analyze data, propose claim, support claim with evidence, prepare individual and group arguments and negotiate on the ideas through small group and whole class discussions (Hand & Keys, 1999). Kingir, Geban and Gunel (2013) claimed that students taught by traditional instruction cannot make further or deep explanation because these students learn through memorization. On the other hand, ABI deals with process of construction of knowledge instead of knowledge itself. For example, when students discuss deficiencies and strong sides of different ideas in negotiations phases, they reach consensus. While reaching consensus, they construct the knowledge. This brings about meaningful knowledge and further explanation.

Likewise, Kingir et al. (2013) claimed that ABI aligns with conceptual change approach. For example; when students reach conflicting data obtained from experiments (e.g. data collection), they are confronted and they become suspicious to their prior knowledge that might include misconceptions or false knowledge. After that students can discuss on conflicting data in their small group and whole group discussion. Students are expected to reach negotiations in these discussions. When students' discussions reach negotiations, students learn from each other. As a result, their initial ideas change and their content knowledge increases. Moreover, ABI includes student template. By using this template, students connect arguments' elements (e.g., claim and data), and link their old and new knowledge that bring about improvement on their content knowledge.

Similarly, Akkus, Gunel and Hand (2007) claimed that prompts of student template contribute on deeper understandings. Accordingly, some students come to the class with their personal knowledge that is unscientific. However, they are supposed to learn scientific knowledge. Learning this scientific knowledge is not easy for students. In this point, student template can be seen as scaffolding. This scaffolding assists students to make reasoning on their work, thinking and discussion. By this way, student template facilitates students learning. For example; fourth question of student template asks students to write their claim and fifth question asks students to write their evidences. When students answer these two questions, they automatically construct their arguments. Without this scaffolding, they may not construct their arguments. Constructing their argument can be seen as a step to reach scientific knowledge. Therefore, it can be claimed that student template is a bridge that connect students' personal knowledge and scientific knowledge.

Similarly, Chen et al.'s (2016) views also support the notion that ABI supports content knowledge. While explaining ABI's contribution to content knowledge, Chen et al. (2016) pointed out the interaction of students' written arguments and whole argumentation process. Accordingly, while using SWH student template, students write their initial arguments. After that they engage in argumentation process based on their initial arguments written in SWH template. By engaging in argumentation process, they reach social negotiation. After, they reach social negotiation, students write their final arguments to SWH template again. When students compared their initial and final arguments written in SWH, they understand how much they learnt in that topic. The differences between initial and final arguments are also evidence that ABI increases content knowledge.

As it is seen, both argumentation and argument-based inquiry are supposed to increase students' content knowledge depending on theoretical explanations of science educators. The next part informs how argumentation affected students' content knowledge in practice.

### 2.2.2. Studies on Argumentation and Content Knowledge.

In this part, firstly, general argumentation studies examining the impact of argumentation on content knowledge are reviewed. After that, argument based inquiry studies that focused on content knowledge are reported. This parts ends with the summary of previous research regarding the impact of argumentation on content knowledge.

Argumentation studies are not always held with immersion orientation (e.g., ABI), there are some other argumentation studies that adopted other argumentation orientations which are socio-scientific orientation and structure orientation (i.e. socio-scientific context and learning the structure of argument excluding immersion ways) (Cavagnetto, 2010). These studies also focused on argumentation and content knowledge. Argumentation studies which are not held with immersion approach examined either the relation between argumentation and content knowledge (e.g., Sadler & Donnelly, 2006) or impact of argumentation on content knowledge (e.g., Dawson and Venville, 2013).

Accordingly, some of the research seeking relation between argumentation and content knowledge was conducted adopting socio-scientific orientation in international context (Dawson & Venville, 2013; Jönsson, 2016; Sadler & Donnelly; 2006). For example, Dawson and Venville (2013) adopted socio-scientific orientation and examined the relationship between argumentation and content knowledge. This study aimed to increase high school students' argumentation skills, informal reasoning, and content knowledge (CK) in the context of socio-scientific issues (SSI) related with genetics through explicit argumentation teaching. This study was mixed method quasi experimental design because there was no random assignment of participants in experiment group (EG) and control group (CG). While the experiment group included 133 students, control group included 160 students. Teaching lasted nearly 10 weeks/ 40 hours. Explicit argumentation accompanied with CK instruction. Argumentation instruction was not covered in CG. Students also produced written arguments based on scenarios provided by their teachers. Teachers in experimental groups encouraged students to discuss with each other. Teachers in experiment group also explained elements of arguments and gave examples of it, encouraged use of claims and counterarguments, encouraged students' use of written arguments, encouraged students to make justification, asked students to examine source of evidence, bias about evidence, validity and reliability of evidence, asked students to think on counterarguments further. Teachers in experimental group also requested students to reflect on argumentation and asked students why they changed their initial arguments. Because of these reasons experiment group might outperform control group in selected variables which are argumentation skills, informal reasoning, and CK. For example; when students are supported to use counterarguments, examine evidence and make justification, their argumentation skills might increase. Likewise, discussion let students to justify their arguments, and they looked for data that supported their claims. By this way, students learnt their arguments better. Cumulative sharing and development of knowledge provided students to learn the topic all together increasing their content knowledge. Similarly, making reflection on their arguments and thinking alternative points might increase their informal reasoning because students might gain multiple perspectives related with informal reasoning when they criticize their own arguments and focus on alternative ideas. The effect size for argumentation skills was medium, for informal reasoning was large and for content knowledge is medium in favour of experimental group.

In another study conducted with socio-scientific orientation, Jönsson (2016) focused on the relationship between content knowledge and argumentation skills based on 12 years old Swedish children's answer to nation-wide exam including argumentation task and content knowledge task. This study also examined different domains of science like physics, chemistry and biology as context for creating arguments. Although there are many studies in argumentation in high-school level, there is limited study in compulsory years. Thus, studying with 12 years old students' argumentation may inform researchers about the developmental effect on making arguments. In this study, different tasks were provided to students. First task is related with argumentation skills and it is formed by three sub-tasks including the use of CK in arguments, selecting correct sources to be used in argument and interpreting and presenting arguments to others to persuade them by use of different representations (table, graphs etc.). Use of content knowledge was provided through concept cartoons and a fictional character sent sources that can be used in argument through email in order. Students were supposed to separate relevant sources from irrelevant ones. Lastly, a video and text provided information to be used in persuasion of others. This argumentation skill task included open ended items. The relationship between argumentation and content knowledge was statistically significant although these relationships were stronger for biology than physics. Findings also suggested that participants had more difficulty in using content knowledge in argumentation than persuading others using different representations showing that argumentation skills are different constructs. Likewise, these skills were not correlated with each other. In conclusion, the use of content knowledge in backing and justification and students' lack of CK cause that students could not provide valid arguments. Lack of argumentation skills are another factors regarding poor arguments. Thus, both content knowledge and argumentation skills education are necessary for students to make valid arguments.

Sadler and Donnelly (2006) focused on how content knowledge and morality contributes on argumentation quality in high school level in socio-scientific orientation (i.e., gene therapy and cloning issues). Content knowledge (i.e., CK) questions about gene therapy and cloning included multiple choice tests. Quantitative data were collected from 125 high school students whose range between 15-18 ages. A total of 48 students were interviewed. The relationship between variables (content knowledge and moral reasoning were predictor variable, and argumentation level was criterion variable) was analyzed through multiple regression analysis. Argumentation level was analyzed through constant comparative method and inductive coding. First initial patterns were revealed by the analysis of transcriptions of 10 interviews. Second, all data were analyzed and patterns derived from initial analysis and second analysis compared. Quantitative findings suggested that neither CK nor moral consideration significantly contributed on the variations in argumentation quality. Qualitative analysis showed that only few participants used their CK in their arguments (5 over 48). Regarding CK, participants' fictional views caused by media replaced with scientific CK. Likewise, some of the participants reported they had lack of CK. First explanation for the use of CK in argumentation is that students do not use their CK in argumentation based on socio-scientific issues. They use CK in school only when they make scientific explanation. Second explanation is related with contextual knowledge. If students are not familiar to context, they may not reflect their CK in argumentation even they know the CK. However, development of contextual knowledge also requires CK. Third explanation is the reaching threshold level of CK that means students having CK less than threshold level cannot use their CK in arguments. There is also second level threshold CK matching with advance knowledge. Accordingly, students having advanced knowledge like majors in a discipline can use their CK in arguments at most. High school students reach first threshold level, thus they can use their CK in their arguments to some extent. Moreover, threshold level graphic showing the relation between CK and argumentation included hypothetical untested population who were deprived of CK not reaching threshold level CK. This people are not expected to use CK in their arguments. Middle school students may be part of this population, so they may not have enough CK and they cannot use them in their arguments. If we think that CK is learnt in social context as in sociocultural perspective, CK and contextual knowledge become dependent. This can connect second and third explanation for the use of CK in argumentation in socioscientific issues.

In Turkish context, Çetin, Doğan and Kutluca (2014) used socio-scientific orientation to reveal the relation between content knowledge and argumentation. Researchers claimed that CK can be related with argumentation more in topic which are not directly with values (topics not directly refer SSI). This study focused on not only CK and argumentation relationship but also possible sources of this relationship in this case study where 54 pre-service science teachers participated in. Participants got CK test and they were divided in three groups as low achievers (scored lower than M- SD/2), high achievers (scored more than M+SD/2) and middle achievers the ones whose scores were between other two groups. All groups worked including three or four pre-services. After dividing groups based on their CK, three scenarios were provided to assist groups to produce their arguments. To understand the factors shaping arguments semi-structured interviews were held. Before, groups produce their arguments; they were trained for components of arguments proposed by Toulmin. Next, they produced and presented their arguments. Pre-service teachers' arguments were analyzed under 5 levels. Uses of rebuttals including other components were indicator to label arguments in different levels. Arguments were produced as dialogs in groups rather than individual statements. The most qualified arguments were revealed by high and low achievers. Moreover, low achievers mentioned rebuttals more than other groups. Findings were also supported by quantitative data analysis. Thus, researchers claimed that CK is not the good indicator for argumentation. Researchers also reported that popularity of the topic, participants' interest and experience are important in making argumentation. But, they also added that CK might strengthen the justification although it does not affect the whole structure of argument. Argumentation skill is another dimension that needs to take attention because CK does not solely guarantee for good argumentation. Researchers also added that values and beliefs are important factors for argumentation, therefore they may be used interchangeably with CK in argument producing process and so CK loses its significance in argumentation. Next, researchers added that triggering participants' interest is important in argumentation. Finally, this study revealed that interaction between group members are important in argumentation because middle achievers reported that newcomers influenced their arguments' quality, and their standstill period in argumentation increased by decreasing their arguments.

As some researchers adopted socio-scientific orientation in argumentation studies, some researchers used structure orientation. Accordingly, Heng, Surif and Seng (2015) examined the relation between argumentation and content knowledge (i.e., acids and bases topic). A total of 120 high-school students received argumentation test as groups or individuals. Each group included four individuals. Individuals and groups' arguments quality were compared with each other. Quality of arguments was classified as valid (scientifically correct) and invalid (scientifically incorrect). Validity of arguments and its elements were assessed quantitatively according to the results of argumentation test; whereas, argumentation schemes including triplet relationship of representations of scientific content, arguments' elements and validity of content were assessed qualitatively. Interview data, however, were used to get deep information about students' arguments. After students got argumentation test, groups and individuals' arguments were compared in terms of validity, elements of argument, and representation of content in acids and bases topic. Findings showed that group arguments are better than individual arguments and this is evidence that argumentation is both cognitive and social process. Although groups performed more accurate arguments than individuals, they also produced misconceptions in their arguments and their confidence was higher than individuals regarding their misconceptions. The results also included some specific findings. Accordingly, findings suggested that individuals did not discuss with their peers in producing their arguments unlike groups. Both group and individual arguments were mostly invalid and included misconceptions. They preferred to use their beliefs and opinion instead of data when producing their arguments. However, groups' performance was better than individuals in terms of validity of arguments because there was knowledge pool in groups. Different ideas and cognitive strengths were integrated. Students also had chance to adjust their ideas after taking feedback from their peers. Findings also suggested that both groups and individuals could not release the elements of arguments. However, groups proposed more elements of arguments than individuals because groups explained their claims to friends through justifications cooperatively. On the other hand, individuals could not gain these justifications. Participants also produced simple arguments that did not include rebuttals and qualifiers. Groups produced more complex and valid arguments including macro and micro representation of the content, but individuals' simple arguments included only claim and data. Warrants were used as macro aspect of content in individuals' arguments. Misconceptions also dominated the individuals' arguments. Why participants could not make justification for their arguments can be related with the absence of link between macro and micro representation.

In another study adopting structure orientation, Aydeniz et al. (2012) designed an experimental research to understand impact of argumentation on content knowledge in Turkish context. Researchers focused on impact of argumentation on CK about properties and behaviours of gases. A total of 108 undergraduate pre-service chemistry teachers participated in this experimental study that lasted two weeks. Control group got 6 hours lecturing and 2 hours of problem solving activity. On the other hand, experiment group got 6 hours lecturing and 2 hours of explicit argumentation instruction followed it. In experimental group, argumentation instruction included written argumentation and verbal argumentation. Students answered two tiered 5 questions. They investigated validity of claims in their written arguments, researchers explicitly addressed structures of arguments in this process lasting about 1 hour. Then, students engaged in verbal argumentation in same questions as groups, this process also lasted in 1 hour. Regarding CK, pre and post-test including 10 questions were administered. In these questions, first, students were expected to select an answer, and then they selected the reason of their answer. Data was analyzed based on independent sample t-tests, and pair sample t-tests. Findings showed that experiment group's CK improved more than control group's CK. Moreover, experiment group eliminated more misconceptions than control group did. Next, experiment group eliminated their misconceptions after argumentation instruction showing that argumentation is effective strategy to abandon misconceptions.

Similar to studies adopting structure and socio-scientific orientation, immersion orientation was commonly used by international researchers to understand the impact of argumentation on content knowledge (e.g., Cronje et al., 2013). ABI is an example of immersion approach and there are many studies using ABI approach to improve undergraduate students, pre-service teachers, high school students, middle school students, students with disabilities and primary school students' content knowledge.

Accordingly, Greenbowe et al., (2007) examined the effectiveness of argument-based inquiry approach over traditional laboratory format on chemical equilibrium equations. This study included 52 science and engineering major undergraduate students. Students' understandings about chemical equilibrium were assessed by use of lecture exam question (i.e., a two tier question that requires selecting and alternative and providing a reason for this selection) and laboratory practical exam task. Inter-rater agreement among scorers was found 0.91. ANCOVA results showed that experiment group using argument-based inquiry approach statistically outperformed control group following traditional laboratory format regarding identifying equilibrium condition and eliminating students' confusions about concentrations change in equilibrium. Researchers claimed that students in control group were better in writing equilibrium equation; however, this was not a good indicator of conceptual understanding. On the other hand, students in argument-based inquiry approach identified the conditions for equilibrium and explained it better which is a good indicator for conceptual understanding. For example; experiment group explained the changing concentrations and its effect on equilibrium and they could connect reactants and products leading better conceptual understandings.

In another study held with undergraduate students, Greenbowe, Poock, Burke and Hand (2007) used argument-based inquiry approach to improve 78 first year college students' general chemistry understandings including measurement, atoms, molecules, elements, thermochemistry. Participants' prior chemistry knowledge was assessed by use of American Chemical Society California Diagnostic Test. Study lasted two semesters and study was implemented by teaching assistants. Researchers observed implementation of instruction in both semesters. Researchers reported that some implementers' use of ABI was better than others. The ones who used ABI in high quality were reported as high implementer and their classes were labelled as high ABI implementation class. On the other hand, other implementers who did not conduct ABI were reported as low implementer and their classes were labelled as low ABI implementation class. In second semester, some of the students changed their classes. Therefore, four different groups of students were created for two semesters. These groups were H-H group which means the groups of students who got high implementation in both semesters, L-L group which means the groups of students who received low implementation in both semester, H-L group meaning the groups taking high quality implementation in first semester and L-H group who took low quality of implementation in first semester and high quality of implementation in second semester. Students' total semester points obtained from homework, lecture assignment, quizzes, and exams were used to understand effect of argumentation based inquiry on their content knowledge. Findings suggested that there was a statistically significant difference between groups' content knowledge at the end of the first semester. Accordingly, students assigned in H-H group performed better than L-L and L-H group for the first semester results. Besides, H-H group students statistically performed better than L-L groups. Students' achievement ranking depending on SWH implementation level was H-H, H-L, L-H, and L-L respectively.

To sum up, this study provides evidence that students' conceptual understandings improve more if the quality of argumentation-based inquiry instruction increases.

Likewise, Shamuganathan and Karpudewan (2017) investigated the effect of argument based inquiry approach on pre-university students' environmental literacy, comprised of pro-environmental attitudes, beliefs, knowledge and behavior. Research design included experiment group in which 120 students participated in and control group in which 90 students (18 years old) involved. Treatment in Experiment group included green chemistry curriculum with SWH approach, on the other hand, control group's treatment included only green chemistry curriculum. Green chemistry is based on protecting natural environment by using non-dangerous resources. Green chemistry education aims sustainable development and is expected to increase students' environmental literacy. In their green chemistry education, for example, students can focus on how factories work without emitting dangerous pollutants. In this process, they are expected to conduct experiment and increase knowledge about green chemistry. What is more, green chemistry curriculum includes laboratory experiments, classroom activities and discussion of socio-scientific issues. Data were collected by use of Questionnaire on Environmental Literacy as pre-test and post-test and data collection process lasted 8 weeks. Data were analyzed by use of one-way analysis of covariance and pre-test scores were used as covariate. Type of treatment (e.g. SWH with green chemistry) was independent variable, and components of environmental literacy which are knowledge, attitude, beliefs, and behavior were dependent variables. Findings of the study showed that experiment group outperformed control group in all four components of environmental literacy. Experiment group became aware of the environmental issues through argumentation and they felt themselves as part of environment and they wanted to protect environment. By this way, their attitude might increase. Likewise, experiment group students provided more detailed information about the issues including deforestation, pollution and hazardous gases. Their detailed explanation might improve their beliefs towards environment too. Moreover, experiment group students focused on conservation of energy and recycling waste. In these speeches, students showed tendency to change their behavior. By this way, experiment group might outperform control group in behavior variable. Similarly, their collaborative and individual participation in knowledge construction process in ABI and taking responsibility for their learning might increase experiment group content knowledge further. For example; experiment group focused on more content than control group such as cleaning of rivers and lakes, ecosystem, food chain, chain reaction in aquatic world, fertilizers, and pesticides. In conclusion, this study showed that SWH with green chemistry education is more beneficial than green chemistry education alone.

In contrast, all ABI research was not as successful as the above mentioned studies to increase content knowledge regarding outperformance of experiment group over control group (e.g., Cronje et al., 2013; Rudd, Greenbowe, Hand & Legg, 2001). For example; Cronje et al. (2013) investigated the effect of argument-based inquiry on 98 undergraduate students' logical conclusions and use of evidence to support conclusions. In the context of the study, there was a biology laboratory course and questions about ecology and evolution were asked to the students. Then, students were supposed to collect data and statistically analyzed them. After that, students were expected to prepare a report that answer research questions. Students' papers were coded ranging from 0 to 6. Low scores were below 3.5 points and high scores were above 3.5. Inter rater agreement was 0.94. T-test analysis suggested that overall scores of argument-based inquiry group were not statistically different from control group. Chi square analysis, however, showed that argument-based inquiry group significantly more likely have high score (above 3.5) than control group. Researchers claimed that argument-based inquiry assists students to connect evidence to claim because of the scaffolding of SWH and students can reflect these arguments in written reports. On the other hand, researchers thought that argument-based inquiry group had to investigate pre-determined hypothesis and use pre-determined data collection procedure. These adjustments might decrease the effect of argument based inquiry approach. Students might think that there is one right answer to be reached as response to pre-determined hypothesis. Likewise, students might think that this research is not their own research that decreases their interest to the research. Researchers also claimed that they used argument-based inquiry in only one activity and this approach should be used in whole semester or more than one activity to better understand the effect of argument-based inquiry. Moreover, researchers claimed that they did not strictly follow the argument-based inquiry sequence although they had done all steps. Researchers added that if they had followed the steps strictly as it was suggested, the effect of argument-based inquiry might increase.

Similar to Cronje and colleagues, Rudd et al., (2001) examined the impact of argument-based inquiry approach on undergraduate chemistry students' conceptual understandings about physical equilibrium topic. A total of 80 students participated in the study. Four different groups were formed and two of them were control group (n=34) who completed standard laboratory format and other two groups (n=46) were assigned in experiment group who enrolled in argument-based inquiry instruction. Students' baseline content knowledge was assessed by use of California Chemistry Diagnostic Exam and ANOVA results showed that there was no statistical significant difference between experiment and control group. Students' response to a physical equilibrium problem was used as post-test and students' answers were coded quantitatively by two raters. Inter rater reliability was calculated as .88 for written explanation and 0.94 for equation usage. Researchers also used a survey in which students were asked to compare SWH format and standard laboratory format to reveal the possible factors contributing on students' conceptual understandings. Median scores showed that students' in argument-based inquiry group had higher response level than students in control group. Chi square results also showed that type of format (i.e., SWH format vs. standard lab format) was correlated to students' written performance. Although there was an effect of type of format on students' written explanation, chi square results showed that there was no significant relation between experiment and control groups regarding equation usage. Researchers thought that equation usage was related with symbolic relationship rather than conceptual understanding, thus groups did not differ. On the other hand, two factor ANCOVA was conducted to see the effect of treatment and effect of implementer on student performance. ANCOVA results showed that there was no main and interaction effect which means there was no statistically significant difference between different groups' student performance in physical equilibrium topic, although descriptive results showed improvement in experiment groups. In survey results, students thought that using SWH format is more efficient than standard laboratory format because it reduced the time on task for students. Moreover, students thought that SWH format is better for them because it provides better understandings and more thinking. Likewise, students were satisfied with directions given in argument-based inquiry.

ABI studies were not limited with undergraduate students. For example; Hohenshell and Hand (2006) examined the impact of argument-based inquiry approach on students' learning about cell topic. The study lasted 7 weeks and 91 Grade 9 students participated in the study. In this process, students completed six laboratory activities and written laboratory activities. While experiment group who got argument based inquiry approach completed SWH lab reports, control group who followed traditional lab format completed formal lab report that starts with hypothesis, goes on recording materials, procedures, results, discussion, and conclusion. Pre-test results showed that there was no statistical significant difference between groups regarding their content knowledge. Content knowledge tests included both lower order recall questions and higher order conceptual knowledge. Students were also expected to write summary reports. While control group's report audience was teacher, and one of the SWH group's audience was teacher, other SWH group's summary report audience was students. Data was analyzed by use of analysis of covariance (ANCOVA) in which gender and treatment were categorical variables, pre-test scores were covariates and conceptual and recall question scores were dependent variables. Findings suggested that experiment and control group did not differ in content knowledge when groups completed lab reports. However, this analysis was done before participants wrote summary reports to their audience. On the other hand, there was a statistical difference between groups after they completed their summary reports. Findings suggested that SWH group who wrote summary report to the teacher performed better than control group who wrote summary report to teacher regarding conceptual understanding. Likewise, SWH group performed better than control group with respect to recall questions. Researcher found no statistical difference between SWH groups in terms of audiences (teacher vs. students) for recall questions. Regarding conceptual questions, females who wrote summary to peers outscored males who wrote summary to peers. On the other hand, there was no difference between females writing to their peers and females writing to their teachers. Moreover, SWH group outperformed control group in conceptual questions.

ABI research was also held in middle school level to see its effectiveness on students' content knowledge. Accordingly, Hand et al., (2004) focused on the effect of ABI instruction on students' content knowledge and metacognition. Research questions were: "If ABI is efficient? If ABI is assisted by writing activity (textbook writing), is it more efficient?, What do students attribute to their own content knowledge development?" There were three groups and two of them were treatment groups in this study. The topic was cell taught in seventh grade. A total of 93 seventh grade students, five classes, participated in the study. Study lasted eight weeks each of which included five 45 minutes periods. Researchers also interviewed with 12 students to reveal their ideas about ABI and their attribution of learning. Pre-post content knowledge interviews were conducted by using multiple choice and constructed response questions. Groups decided their own research questions. Only control group instruction was based on step by step instruction. ANCOVA was used to understand differences among three groups. Interviews were conducted with students who got ABI instruction. Cronbach alpha score was found as 0.61 for pre-test and 0.89 for post-test. Treatment group who took both ABI and wrote textbook explanation outperformed only ABI taking group and control group in conceptual questions post-test. Both treatment groups outperformed control group in post-test multiple choice questions according to ANCOVA results. In qualitative results, students thought that they understand better because they asked their own research questions, participated in discussions, linked the difference concepts, and wrote their ideas. Writing their own research questions increased their interest and they became more curious to complete their research. Students also reported that they learn more because in discussion, they both learn by listening others' ideas and by sharing their ideas and getting comments on their shared ideas. Group discussion assisted students because they could produce group claim. Moreover, group discussion let students to share cognitive load for generating claims. Students also found ABI efficient because they thought that ABI assisted students to understand how they learnt. Researchers claim that although previous studies did not find a difference between groups in multiple choice questions, this study found a difference between groups. SWH increased ownership, responsibility, and inquiry. When students increased their responsibility by seeking an answer for their own research question, they increased their content knowledge.

In another study; Hand et al., (2016) focused on elementary students' development on science and language, based on different quality of ABI implementation and socio-economic status. Iowa Test of Basic Skills was used to understand students' growth. Study lasted three years. A total of 780 students participated. First year, 31 teachers; second year, 32 teachers; and third year, 32 teachers participated in the study. Quantitative data was analyzed by use of ANCOVA to understand students' growth in science and language. Findings showed that level of implementation affected students' science growth. Likewise, findings showed that students' science growth was different between low SES and medium SES in low and medium implementation with small effect size, but in high level implementation, there was no difference between students' science score growth.

In another study examining the impact of collaboratively writing letters on fourth grade students' conceptual understanding about force and motion unit; Chen, Hand, and McDowell (2013) used quasi experimental pre-test post-test research design. In this study, 835 fourth grade students wrote letter to 416 $11^{th}$ grade students about force and motion. The teaching lasted 8 weeks and fourth grade students wrote three different letters about topic throughout the process. After fourth grade students wrote their letters, $11^{th}$ grade students provided feedback and asked questions to the fourth grade students. The focus of students' writing was claim, evidence and question component of arguments. Findings of the study showed that students who got treatment performed better than other students who did not complete collaboratively writing letters. Moreover; female, gifted and disadvantageous students benefitted from the study further. Throughout the process, $11^{th}$ grade students supported fourth grade students' learning. When 11th grade students asked for clarification and explanation, fourth grade students learnt more. When students correctly used multi-modal representation and connected claim and evidence, fourth grade students became more successful.

ABI research was also conducted with students having disabilities. Taylor et al. (2018) examined the effect of SWH instruction on students with disabilities science achievement. Researchers claimed that students with disabilities science performances are worse than students without disabilities and this gap increases over time. Therefore, there is a need to close this gap between students with disabilities and other students. Thus, SWH approach can be used as a tool to assist students with disabilities content knowledge. Moreover, literature review suggested that when inquiry is accompanied with more support, students with disabilities learn better, and scaffoldings of SWH can be helpful for students with disabilities. There were 407 third, fourth and fifth grade level students in the study and 208 students were assigned into treatment group who took SWH teaching and 199 students were assigned into control group. On the other hand, control group instruction included textbook based instruction and lecturing. The Iowa Test of Basic Skills was used as pre and post-test to measure participants' content knowledge including reading, comprehension, math, science, scientific inquiry, life science, physical science etc. T-test results showed that there was a statistically significant improvement on students' content knowledge from pre-test to post-test for both experiment and control groups. Regarding treatment effect, experiment group outperformed control group statistically. Moreover, there was a moderate level effect size between experiment and control groups. In conclusion, results showed that SWH is a useful approach to improve students with disabilities content knowledge.

ABI instruction has been also popular among Turkish science education researchers. In Turkey, researchers focused mainly on the impact of ABI instruction on pre-service teachers, high school students, and middle school students. Regarding studies held with pre-service teachers, Demirbag and Gunel (2014) integrated ABI with multi-modal representation. According to this instruction; ABI was enriched by use of multi-modal representations which are the tools such as picture, text, diagram, or mathematical expression. Researchers examined impact of this treatment on pre-service teachers' content knowledge, argumentation, and writing skills. Totally, 119 pre-service participated in study and four classes formed. Two of the classes got ABI instruction, and two other classes applied ABI with multi-modal representations. Data were collected by use of exam scores and written reports. ANOVA results showed that there was no difference between groups' prior content knowledge. On the other hand, ANOVA results showed that ABI with multimodal group outscored the only ABI group based on the analysis of students' exam scores and written report scores. Moreover, the group taking ABI with multimodal representation performed better than the group who took only ABI in terms of argument quality scores. Pearson correlation coefficient between multimodal representation coherence and argument quality was found .65 indicating strong relationship.

Similar to Demirbag and Gunel (2014), Erkol, Kışoğlu, and Gül (2017) examined the effect of ABI instruction on pre-service teachers' physics content (i.e., matter, density, simple machines, force and pressure) knowledge and their attitude towards science. A total of 52 pre-service teachers participated in the study. Half of the students formed experiment group and the other half formed control group. While experiment group prepared SWH reports, control group prepared traditional science laboratory report. Descriptive statistics and ANCOVA was used to test the effectiveness of ABI on content knowledge and attitude towards science. Findings suggested that experiment group outperformed control group in science content, but groups did not differ in terms of attitude towards science. Researchers added that the reason why experiment group did not differ from control group might be their approach. Accordingly, other aspects of ABI such as negotiating on the ideas through small group apart from using SWH report were not emphasized in experiment group. Therefore, only focusing on one aspect of ABI might be insufficient to increase students' attitude towards science.

Recently, Yaman (2019) examined how students construct their multi-level representation in ABI instruction. Data was collected from seventy-seven pre-service elementary teachers in physics and chemistry topics. Student template of SWH was used as data collection tool and students generated 156 laboratory reports. Analysis of students’ lab reports revealed three assertions regarding multi-level representations. First assertion claims that students used symbolic level representations more than macroscopic and microscopic level in their lab reports. Moreover, they connected symbolic level and macroscopic level more than other possible connections. Second assertion shows that symbolic level play key role in students’ writing. Symbols were connected to macroscopic level such as colour change and to microscopic level such as bonding of molecules. Third claim is related to SWH. Accordingly, students connected representations in evidence and reflection parts of SWH. Likewise, students used representations in data part of SWH at most.

Regarding high school level, in Turkish context, Kingir and her colleagues conducted ABI research to reveal its effect on high school students’ content knowledge. One of these studies examined the extent to which argumentation-based inquiry approach is effective for students having different academic achievement level (Kingir, Geban & Gunel, 2012). The study was conducted in chemical change and mixtures topics. This study was conducted in 9th grade level (15-17 years) and 122 students participated in the study. Prior to the study, students were divided in three parts for their achievement levels based on their previous semester grades. These academic levels were low (mean-SD/2>low), medium (mean-SD/2<medium<mean +SD/2), and high (mean +SD/2<high). The teachers applied traditional lessons for control group, and SWH for treatment group. Traditional lessons mainly included lecturing and discussion methods. Students also conducted experiments which are based on verification and teachers controlled this process in control group. The study lasted 10-week period. Students’ conceptual understanding was measured by content knowledge test that included 22 multiple choice questions. ANCOVA was used to analyze effectiveness of the ABI treatment, and students’ pre-test scores were used as covariates. Findings suggested that there was a statistical significant difference on post-test scores of content knowledge between different treatment groups in favour of experiment group. The effect size was large. Findings also suggested that low, medium and high achievers in experiment group performed similar to each other showing a decrease in achievement gap. There was no statistical difference between groups in terms of academic levels. However, different academic

levels' performance was different where high achievers developed themselves most and low achievers developed themselves the least in control group. While low achievers of experiment group outperformed low achievers of control group, and medium achievement level in experiment group outscored medium achievement level in control group, there was no statistically significant difference between high achievers of different treatments.

Another research by Kingir et al., (2013) examined whether Argument-based Inquiry affect 9$^{th}$ grade students' CK in chemical change and mixture topics. A total of 122 students participated in the study. There was no random assignment in this quasi-experimental study because students were already enrolled in their class and they could not be re-assigned in classes for the aims of the study. Some variables that might affect the results were controlled by evidence derived from descriptive statistics (i.e., mean score of variable for both groups). These controlled variables were previous chemistry scores, age, gender, and SES. Data were collected by using CK was two tiers where first part included multiple choice questions and second part included reasons or open ended explanation. Further information was obtained from 21 students through interviews. Study lasted 10 weeks and each week included two hours. Control group also conducted lessons in lab in order not to decrease their attitude towards science. Moreover, this group conducted textbook based experiments which are recipe-type. On the other hand, experiment group followed ABI instruction. Findings showed that ABI group's CK was better than traditional group according to ANCOVA results. Medium effect size was observed. Pre-tests were covariate in this analysis and there was difference between pre-test scores in favour of experiment group, but this effect was eliminated by use of covariate in findings. Control group were found to have more misconceptions than experiment group. Analysis of interview data obtained from both experiment group and control group showed that experiment group defined concepts better than control group. Both groups provided similar examples of the concepts, such as physical change, chemical change, mixtures, and solutions. Regarding the relationships between concepts such as mixture and solution; treatment group performed better. Researchers discussed the reasons why experiment group performed better than control group considering the embeddedness of conceptual change approach in ABI instruction. Accordingly, students deal with their prior knowledge, and solve the conflict between prior and new knowledge, and apply new concepts in new situations. By this way, while experiment group improved their content knowledge, control group who followed recipe type laboratory work could not develop their understanding.

Gunel, Kingir and Geban (2012) also conducted studies including middle school students. They examined student and teacher questions in ABI treatment and general patterns of questions and their relations with argument construction were investigated. Three teachers and their 146 sixth grade students participated in the study. Data were obtained from video records. Findings suggested that teachers asked more questions than students did. The level of questions that teachers asked included high level and low level questions. Teachers differed in starting and sustaining negotiations. When teachers asked low level questions, there were little negotiations. On the other hand, higher level questions accompanied with more negotiations. Moreover, follow up questions fed negotiations. Moreover, how ABI was implemented affected negotiations. Accordingly, when teachers controlled the class and had all ideas, they did not let so many students talk. Moreover, teachers judged and evaluated students' response when they were controller. On the other hand, teachers gave importance to students' talk when they acted as guide. When teachers acted as facilitator and guide in ABI classes, they asked more high level and follow up questions leading more student talk.

Likewise, Yeşilçağ-Hasançebi and Günel (2013) examined the impact of argument-based inquiry approach on 55 eighth grade level students learning about properties and structure of matter topic. The school of students was located in a disadvantaged social economic area of Erzurum. While traditional teaching was applied to the control group, experiment group got argument based inquiry approach. Content knowledge test included multiple choice items and open ended questions. Findings suggested experiment group performed better than control group both in multiple choice and open ended questions and the effect size was large in favour of treatment group. Researchers claimed that ABI approach assists disadvantageous students' learning. Moreover, researchers claimed that developing argumentation skills of students throughout the process facilitates their science learning. Lastly, researchers claimed that there was a positive correlation between quality of students' SWH template and their learning. That means when students' arguments are supported with evidence and there is coherence between research question, claim and evidence in students' arguments, they learn the science topics better.

### 2.2.3. Summary of Argumentation Studies and Content Knowledge.

To sum up, related literature showed that the relation between argumentation practice and content knowledge is uncertain if researchers do not adopt immersion orientation (e.g., ABI). For example; Sadler and Donnelly (2006) claimed that content knowledge did not significantly contributed on variations in argument quality. On the other hand, other researchers found either positive relation between content knowledge and argumentation (Heng et al., 2015; Jönsson 2016) or positive impact of argumentation treatment on students' content knowledge (Dawson & Venville, 2013). Unclear relationship between content knowledge and argumentation exist in Turkish literature in which researchers did not adopt immersion approach. While Aydeniz et al. (2012) reported that argumentation has impact on students' content knowledge; Cetin et al. (2014) claimed content knowledge is not a good indicator for argumentation. On the other hand, argumentation impact on content knowledge is clearer when researchers adopted immersion orientation. Only two of the ABI researches showed no significant difference between experiment and control group regarding content knowledge (Cronje et al., 2013; Rudd et al., 2001). On the other hand, rest of the reviewed studies showed that experiment group outperformed control group regarding content knowledge. Moreover, all the studies adopting immersion approach and using ABI instruction in Turkish context showed that experiment groups statistically significantly performed better than control group. These practical findings support theoretical expectations that argumentation feeds content knowledge.

#### 2.2.4. Reflections from Literature to Current Study regarding Content Knowledge and Argumentation.

Next, how previous literature contributed to current research about argumentation and content knowledge is explained. Mainly research design of this study was affected by previous research when participants' content knowledge was aimed to increase thorough argument based inquiry. Accordingly, Chin and Osbourne (2010) claimed that scientific knowledge is constructed when alternative perspectives are explained. Therefore, while lesson plans were being prepared, alternative points were considered. For example, in week 4, students were asked to select best sources of energy (e.g., solar energy), and there were many alternatives. Likewise, there were different advantages and disadvantages for each sources of energy. These advantages and disadvantages were used as starting points of alternative ideas in lesson plans. Likewise, Aydeniz et al. (2012) claimed that peer questioning provides deeper understanding. Therefore, students were encouraged to ask questions to other groups throughout the study. Students were at the centre of teaching and they were active participants. Similarly, Kingir et al. (2013) claimed that SWH student template provides that students easily connect argument component each other. Moreover, students can connect prior knowledge to new topic by use of SWH template, so SWH contributes on learning. Due to fact that SWH is scaffolding for learning, SWH student template was used actively throughout the study. By this way, it is possible that SWH increased students' content knowledge. Moreover, Cronje et al. (2013) and Hand et al. (2004) warned researchers about some factors that decrease the effect of ABI instruction. Accordingly, pre-determined hypothesis and pre-determined data collection procedure decrease the effect of ABI on students' learning. When pre-determined hypothesis and data collection is provided to students, students might think that there is one right answer in science and research is not their own research, so their interest to ABI decreases, they engage in argumentation less, and their content knowledge does not increase. Because of these reasons, when lesson plans were prepared, pre-determined hypothesis and data collection procedures were not used, whenever it is possible in this study. Likewise, Hand et al. (2004) claimed that students learn from each other when they discuss with their peers in line with social constructivism. Therefore, they should listen to each other before discussing with others. In line with this, some pro-active rules were explained to the students at the beginning of the study such as listening to other students when they talk and respecting other students' ideas. One other contribution of the literature was about teacher actions. Accordingly, Günel et al. (2012) reported that when teachers ask high level questions (e.g., how, why questions), there are more discussions in class leading to better learning. Likewise, follow up questions feed discussion. This report was considered when teacher training was prepared. Therefore, in teacher training, teacher was asked to use high level questions and follow up questions in teaching. Moreover, these questions were embedded in lesson plans too. Likewise, Günel et al. (2012) reported that when teacher do not judge students' ideas, and teacher do not dominate the class discussion, ABI's efficacy increases. In line with this, teacher was asked not to judge students' ideas and not to dominate class in teacher training part. Moreover, when lesson plans were being prepared, teacher role was determined as facilitator in line with Günel et al.'s (2012) suggestions.

According to vision 1, scientific enterprise is at the centre of scientific literacy (Roberts, 2007), and scientific literacy aligns with argumentation. Content knowledge that is one cognitive variable related with scientific literacy and its relation with argumentation including ABI was reviewed in this part. Similarly, scientific talking and scientific writing are two aspects of scientific literacy. Through argumentation, students read different sources, conduct experiments, make critics and engage in scientific debates. By doing argumentation like scientists, students learn scientific talking and writing (Jimenez-Aleixandre, & Erduran, 2008). While they engage in scientific practice via scientific talking and writing, their epistemological beliefs which are about construction and characteristics of scientific knowledge may also change. In line with this, previous studies examining the relation between students' epistemological beliefs and argumentation including ABI are reviewed in next part.

## 2.3. Argumentation and Epistemological Beliefs

After review of argumentation studies that deal with content knowledge, this part focuses on argumentation studies including its relationship with epistemological beliefs. Firstly, epistemology and theoretical consistency between argumentation and epistemological beliefs are explained in this part. Next, practical work regarding argumentation and epistemological beliefs including nature of science are presented. At the end, summary of the studies about argumentation and epistemological beliefs are given.

In 1990, Schommer defined epistemological beliefs as multiple distinct beliefs which are developed independent from each other. These distinct epistemological beliefs are stability, structure, source, speed of acquisition, and control of acquisition. Then, using factor analysis, Schraw, Bendixen, and Dunkle (2002) found five dimensions of epistemological beliefs which are stability, structure, source, speed, and control. On the other hand, Hofer and Pintrich (1997) claimed that speed and control dimensions are not related with knowledge and knowing. These two dimensions are about learning.

According to Hofer and Pintrich (1997), epistemological beliefs include four dimensions which are certainty of knowledge (stability), simplicity of knowledge (structure), source of knowing (authority), and justification for knowing (evaluation of knowledge claims). However, all these dimensions were obtained from the studies in which adults participated. Therefore, these dimensions did not represent the younger people epistemological beliefs (Conley et al., 2004). In line with this Elder (2002) proposed different dimensions of epistemological beliefs for elementary students. These dimensions are changing nature of science (stability), coherence of knowledge (structure), source of knowledge (source), and role of experiments (refers to knowledge justification in science).

Following Elder (2002) work, Conley et al. (2004) defined epistemological beliefs as the belief system including source, certainty, development, and justification. According to Conley et al. (2004), sources of knowledge deals with where scientific knowledge comes from. Source can be people who conduct scientific research or outside of the people. Certainty is about whether scientific knowledge is certain or uncertain. Development is about whether scientific knowledge changes or not and justification emphasizes the role of experiment, testing and evidence in science. In this study, Conley et al.'s (2004) conceptualization for epistemological beliefs was used because epistemological beliefs used in their study were used to understand young students' epistemological beliefs rather than adults. Current study also focuses on middle school students' epistemological beliefs and therefore, Conley et al.'s conceptualization is suitable for this study to understand participants' epistemological beliefs. Other epistemological beliefs' conceptualizations (e.g. Hofer & Pintrich, 1997) can be suitable for older students and adults' epistemological beliefs.

According to Appleton (2007), research type has also changed because of the popularity of constructivism. When constructivism dominated the science education, researchers changed their position from positivism to constructivism. In line with this, researchers might expect students to have constructivist ideas rather than positivist ideas in their epistemological beliefs. For example, when Conley et al. (2004) examined students' epistemological beliefs that are about students' ideas related with knowledge and knowing, their position was constructivism. They evaluated the knowledge and knowing from this perspective and so students having constructivist epistemological beliefs were labelled as students having sophisticated beliefs and students having positivist epistemological beliefs were labelled as students having less sophisticated beliefs. In line with this, Epistemological Beliefs Questionnaire (EBQ) developed by Conley et al. (2004) was used in this study in order to assess students' epistemological beliefs. The higher scores obtained from justification and development sub-scales represented sophisticated beliefs as suggested by Conley et al. (2004). On the other hand, Conley et al.'s (2004) questionnaire's source and certainty sub-scales' items were negative items. Students responses to these items reverse coded before analysis. Therefore, students having higher scores in source and certainty sub-scales after reverse coding were accepted as students having sophisticated epistemological beliefs and students having lower source and certainty scores in these sub-scales after reverse coding were accepted as students having less sophisticated beliefs.

### 2.3.1. Theoretical Consistency between Argumentation and Epistemological Beliefs.

Theoretically, the relationship between argumentation and epistemological beliefs has a feedback loop. Before argumentation starts, people epistemological beliefs determine whether they engage in argumentation or not. In fact, McDonald and McRobbie (2012) pointed out that there are three types of epistemological orientations which are absolutist, multiplist, and evaluativist. In absolutist view, knowledge is perceived as certain and fixed. That means there is one-right scientific knowledge in a topic. It does not change and it always exists. However, knowledge is seen as personal opinions and these opinions are equally valid in multiplist view. According to evaluativist view which is hierarchically better than other two views, knowledge is evidence based theories and opinions do not have equal value. Evidence is examined by evaluativist and claims are verified or not. Accordingly, multiplist engages in argumentation less than absolutists and evaluativist. Multiplists also do not consider inconsistencies of their reasoning because they think that every opinion has same validity. On the other hand, evaluationist criticizes others based on available evidence more than other two views. Their inconsistent reasonings are fewer than other views and they deal with various different ideas. On the contrary, absolutists engage in argumentation more than multiplist because seeking correct answer motivates them to engage in argumentation. To sum up, people epistemic view affects their engagement in argumentation process.

Likewise, Liu and Roehrig (2019) claim that people who think that knowledge is relative and not absolute accept that argumentation is central practice in science and these people engage in argumentation practice. In other word, sophisticated beliefs of epistemology let people engage in argumentation practice.  Similarly, Walker and Sampson (2013) emphasize the importance of epistemic beliefs to construct arguments. Walker and Sampson (2013) claimed that students can support their claims, they can produce arguments. However, their argumentation skills are not well developed. For example; they do not reject data. In contrast, they manipulate or trivialize it. They do not recognize the evidence. They rarely use suitable evidence, and focus on the method and details in argumentation. Even though, they know the evidence, they do not use rationale to support claim with evidence. To sum up, students do not know what is important in science. The reason why their scientific arguments are not developed is not their cognitive ability. The reason is their lack of knowledge about how scientific knowledge is produced, and validated. Up to this point, it is obvious that epistemological beliefs are vital to engage in argumentation; therefore, it is better to explain how argumentation contributes on epistemological beliefs to complete the feedback loop. According to Martin and Hand (2009), when students construct and criticize knowledge in social environment, students create scientific community and they make scientific inquiry. When they make scientific inquiry they act as scientists. Through argumentation, science as is taught and science as is practice become closer. In other word, engagement in scientific practice let students to better understand uncertain structure of knowledge, its justification, development and source aspects (i.e., epistemology). Likewise, Sandoval and Millwood (2008) reported that understanding the rules of argumentation represent epistemology of science. Therefore, theoretically, argumentation develops students' epistemology.

According to Sandoval and Millwood (2008), nature of science is related with epistemological beliefs and students' NOS views inform about their epistemological beliefs. As there is a theoretical link between epistemological beliefs and argumentation, nature of science and argumentation are linked too. Accordingly, Khishfe (2014) explains that theoretically, argumentation supports NOS views. This positive relation can be explained based on "counterarguments". In other words, having different claims on same evidence (i.e., counterargument, and argument) aligns with subjectivity aspect of NOS. Counterevidence used for counterarguments supports the idea of science is empirical based. The idea of accepting counterarguments if they are valid and persuasive (i.e., replacement of argument with counterargument) supports that science is tentative. On the other hand, Khishfe

(2012) claims nature of science does not represent epistemological beliefs one hundred percent. Khishfe (2012) reported that NOS shows people views about how knowledge is constructed. NOS specifically shows the beliefs about nature and construction of knowledge that means NOS is related with only one component of epistemological beliefs that is nature of knowledge (i.e., certainty and simplicity of knowledge). However, NOS ignore another aspect of epistemological beliefs which is nature of knowing. To sum up, NOS is just one part of epistemological beliefs. Khishfe added that there are few studies that examine the impact of argumentation on epistemological beliefs.

### 2.3.2. Argumentation Studies on Students Epistemological Beliefs.

After theoretical consistencies between argumentation and epistemological beliefs are explained, research examining the relationship between argumentation and epistemological beliefs including NOS studies are presented in this part.

In her master thesis, Tucel (2016) examined the impact of a 13- week argument-based inquiry approach on 60 $8^{th}$ students' content knowledge, metacognitive skills and scientific epistemological beliefs. Content knowledge test included questions about sound, the relation between living organisms and energy, states of matter and heat, and electricity in daily life. While treatment group conducted science activities using ABI approach, control group engaged in laboratory activities in traditional lab format. The effect of treatment (ABI) on dependent variables (e.g. epistemological beliefs) was analyzed by one way MANOVA Findings suggested that groups did not differ at the beginning of the study with respect to content knowledge, metacognitive skills, and epistemological beliefs. On the other hand, experiment group outperformed control group regarding content knowledge in post-tests and large effect size was reported. Regarding metacognition, researcher reported that experiment group outscored control group in declarative knowledge, planning, information management, monitoring, debugging, and evaluation; however, there was no statistically significant difference between experiment and control group in terms of procedural and conditional knowledge. Next, participants' epistemological beliefs results suggested that there was large effect size and statistically significant difference between groups in favour of experiment group regarding justification and development components of epistemological beliefs. On the other hand, there was no statistically significant difference between groups in terms of certainty and source components. Researcher explained the reason why experiment group students' certainty and source of knowledge beliefs did not differ referring their perspective towards science. Accordingly, students may think that the science that they do in class is not same as the science that scientists do. While students' work results in uncertainty regarding experiments and argumentation and they do not need external authority in their work, they still think that scientific knowledge is certain and it comes from external sources when scientists do science.

As far as studies about the effect of argumentation instruction on NOS are considered, McDonald (2010) examined the effectiveness of argumentation treatment with explicit NOS instruction (i.e., the instruction that NOS tenets are explicitly taught to the students) on 16 pre-service primary teachers' nature of science beliefs using questionnaires, interviews, video-tapes and written artefacts. NOS aspects emphasized in this study were empirical NOS, methods of science, relationship between theories and laws, tentative NOS, inferential and theoretical NOS, subjective NOS, socio-cultural NOS, and creative NOS. In argumentation treatment, participants used argumentation scenarios and two of the scenarios were scientific (Mixtures, Elements, and Compounds, snowmen-heat transfer) and three of them were socio-scientific (i.e. Diet, exercise, and cancer, Cigarette smoking and cancer, fetal tissue transplantation). NOS aspects which were subjective NOS, theory laden NOS, socio-cultural NOS, and creative NOS improved more than other components, but the relation between theory and law improved less. Explicit NOS and argumentation instructions were taught to the teachers prior to the study. In first week, participants observe and explain solubility and behavior of a group of inorganic substances and they engaged in whole class discussion. They also discussed about NOS tenets. They explored principles of separation in second week and content of the discussion was Mixtures, Elements, and Compounds. In forth week, participants engaged in argumentation about first SSI topic that is diet, exercise, and cancer. In fifth week, argumentation topic was in scientific context that is about snowmen where teachers discussed on heat transfer. Teachers engaged in argumentation about the relation of smoking and cancer using the socio-scientific context that is cigarette, smoking and cancer in six weeks. Finally, participants discussed on drugs and medicine in fetal tissue transplantation topic. Participants reported that explicit NOS instruction assisted their understanding about NOS aspects, but they could not provide specific references about. Moreover, participants did not reflect their NOS views when they engaged in argumentation both in scientific and non-scientific contexts. In other word, participants did not explicitly connect nature of science tenets to argumentation and the context they discussed. Participants reported that some factors caused them not to participate in argumentation such as lack of self-confidence, lack of content knowledge and lack of argumentation skills, and participant' ideas about other members of group. For example, some of the group members talked that she did not like personalities of other group members. Regarding participants epistemological reasoning, McDonald claimed that participants might have multiple forms of epistemological reasoning. Accordingly, if participants had sufficient content knowledge, they develop their specific epistemological reasoning in scientific context. For example, participants could not apply their epistemological reasoning in superconductor topic because of their lack of content knowledge. On the other hand, participants reflect their general epistemological reasoning on both scientific and socio-scientific context when they did not have sufficient content knowledge. While participants developed arguments in socio-scientific context, they could not able to develop argument in scientific context. This shows that engaging argumentation in scientific context is more difficult than engaging argumentation in socio-scientific context. Content knowledge as pre-requisite knowledge to engage in scientific argumentation might explain why argumentation in scientific context is more difficult. The students having lack of content knowledge may not able to engage in argumentation in scientific context, to support and to justify their arguments. On the other hand, supplying scientific evidence assisted participants to engage in argumentation in socio-scientific context. By this way, they could justify and support their claims. Likewise, conceptual demands of socio-scientific context are less than scientific context and participants can use their personal experiences and values (e.g., ethical) to construct arguments in socio-scientific context. These differences might make engaging in argumentation easier in socio-scientific context comparing with scientific context.

In another study, Khishfe (2014) included both explicit instruction and SSI topics as context to increase participants' argumentation skills and NOS beliefs. A total of 121 seventh grade students participated in the study. Classes were unit of analysis in this study; therefore researcher reported percentage and frequency for comparison of different treatment groups. Instruction, including 2 kinds of treatment, lasted 8 weeks. In treatment-1, explicit argumentation instruction was provided. In this instruction, components of argument were explicitly taught to students. Strong and weak sides of the arguments were compared and contrasted in class to show quality of arguments. Students were asked to produce counterarguments as well. Treatment 1 (i.e., explicit argumentation and NOS instruction) provided students to analyze argument structures formed in classroom activity. Moreover, teacher in treatment 1 scaffolded students' arguments. On the other hand, treatment 2 (i.e., explicit NOS instruction) included same activities as implemented in treatment 1 except explicit argumentation instruction. Argumentation was analyzed as three separate components which are argument, counterargument and rebuttal. These components were analyzed in three increasing levels. Level 1 included claim without reasoning, level 2 included claims with one reason, and level 3 included at least two reasoning for each different argumentation component. Findings of the study showed that in treatment 1 (i.e., explicit argumentation and NOS instruction), students could not make reasoning at the beginning of the study; however, reasoning increased after instruction. On the other hand, treatment 2 (i.e., explicit NOS instruction) was found to be useful for increasing students' argumentation skills in familiar context. Students in treatment 2, however, could not transfer their developed argumentation skills in unfamiliar context especially for counterarguments and rebuttals. Thus, treatment 1 is more successful than treatment 2 for argumentation development and transferring argumentation skills to unfamiliar contexts. In other word, explicit argumentation instruction can be helpful for transferring argumentation skills into unfamiliar context. Both groups developed their naïve NOS understandings after instruction for both familiar and unfamiliar context. So, they could transfer their informed NOS views from familiar to unfamiliar concept. Groups did not differ in terms of their NOS understandings; however, group who took argumentation included their argumentation experiences in their NOS explanations. Findings also showed that the ones who developed their NOS views also developed their argumentation skills which might show positive relationship between NOS and argumentation.

In another study adopting socio-scientific orientation, Khishfe, Alshaya, BouJaoude, Mansour, and Alrudiyan (2017) examined the relation between 11th grade students' arguments and their nature of science beliefs in four different socio-scientific contexts (global warming, genetically modified food, acid rain and human clothing) and 74 students participated in the study that was conducted in Saudi Arabia. Subjectivity, tentativeness and empirical based aspects of NOS were included in the study. Likewise, argument components which are argument, counterargument and rebuttal were another focus of the study. Data was collected by use of interviews, observations and four open ended scenarios related with controversial issues which are global warming, genetically modified food, acid rain and human clothing. After researchers coded participants' NOS beliefs and arguments as naive, intermediate and informed hierarchically, they sought correlation analysis between NOS beliefs and arguments. For example; when participants did not use any justification, they were labelled as naïve understanding about argumentation (e.g., I support human cloning). Regarding NOS coding, if participants' ideas were not consistent with contemporary NOS views, they were labelled as naïve (e.g., the knowledge about human cloning will not change in the future). Chi square analysis was also conducted to compare NOS aspects (e.g., tentativeness) and argument components (e.g., argument) between different controversial scenarios. Researchers found few and low correlations between argument components and NOS aspects. On the other hand, qualitative data analysis suggested that participants having developed understanding had informed NOS beliefs. Chi square analysis also showed that participants' responses regarding argument components and NOS beliefs were not different between four scenarios. In other word, participants' NOS views and arguments followed similar trends in different scenarios. However, qualitative analysis suggested that participants emphasized different NOS aspects in different scenarios. Researchers thought that different orientations that participants held affected their NOS beliefs and arguments. These orientations were reported as religious orientations, natural orientations, and environmental orientations.

In another study, Mendonça and Justi (2014) focused on the relation between scientific models and argumentation. Accordingly, scientific models are important because they show that scientific knowledge is neither true nor false. This situation also aligns with argumentation because knowledge constructed in argumentation process is not certain too. There are many models to explain scientific phenomena and the most suitable ones are selected through justification and refutations. Evidence in models shows models' power if evidence works in reality (evidence works in field). However, outer factors like persuasion can be important to prefer one model and eliminate others. There is a two directional relation between argumentation and modelling because we use our arguments to construct models, and models assist us to justify our claims as theoretical constructs. When we test models, we evaluate the evidence that model is constructed on which through argumentation. Construction of models through testing shows that models are not copy of reality. Thirty-eight high school students (16-18 ages) participated in the study. TAP does not assist analysing arguments to their components because it is difficult to distinguish the components. Thus, researchers' analysis approach was comprehensive for arguments. Modelling activities were produced in chemical bonding, ionic bonding, and intermolecular bonding. Implementation lasted 9 hours and it was video-recorded. Students conducted experiments before they made argumentation. Class was separated in 6 groups, but data were collected from 2 groups whose data were available. During analysis, researchers made explicit the models students proposed, the arguments about models, and the ideas about the production, explanation and use of models. Then, instrument was prepared. Findings showed that instrument had three hierarchical levels. Level-1 included a claim with single type of justification which can be supported by either empirical evidence, theoretical evidence, or representational evidence. Level-2 included more than one type of justification without persuasion. Level-3 included more than one type of justification which aims to persuade others about the validity of selected model. Level-1 and level-2 was based on the understanding the phenomena by constructing model. In other word, the aim was sensational goal. However, level-3 was based on the persuasion of others. Level-3 shows that people understand the model, test the evidence of model, and prefer this model among others based on testing result. The goal of this level was "persuasion". Level-1 of this study was different from other research's level-1 because level-1 was justified in this study unlike other research. Students started with level-1. Because of students of first group did not understand the phenomena, they could not use empirical evidence. They just used theoretical evidence which was found in their prior knowledge. Second group's model was worse because they were both in level-1 and their prior knowledge was deficient. So, they just used representational evidence. Because of they had no prior knowledge; they could not use theoretical evidence. Then, students reached to level-2. Students integrated theoretical and empirical evidence and reached a model that was consistent with scientific explanation. Next, students tested the model they reached in level-2. The model reached at the end of the level-2 was better than other models because it fit the observation. Students started to use this model to persuade others and they reached level-3. In this level, students not only tried to persuade others about validity of final model, they also explained why previous models were not valid. Validity of model made it more persuasive. Activities used in this study contributed on students' level of argumentation in model based context.

Previous research on ABI also suggested that students' use of data type (i.e., self-collected primary data or second hand data that is obtained from somewhere else) can be related with students' epistemological views. For example; Kolsto and Ratcliffe (2007) reported that students directly use data coming from expert in their arguments. This is related with their naïve epistemological beliefs because they think that expert opinion is perfect and objective, they do not criticize expert opinion and they directly accept them. However, if science is closed to discussion as in the case, debaters discuss based on non-scientific things. On the other hand, some students think that scientists can be criticized and these students want to understand scientists' claims more. If students' epistemological beliefs are improved, their suspicion on claims improves too. Therefore, it is meaningful to report ABI research that focused on the type of data that students use in this part.

According to Wallace (2004), students are expected to emphasize similarities and differences between different information such as textbook and experiment results. In ABI, there are two different kinds of sources. First is first-hand observation which is single, personal and interpreted by theory. Second are textbook statements which are body of knowledge validated by scientific community. These two sources' discourses are also different. In hand on activities, students use informal language; however, ABI makes it more scientific because it adds claims and evidence to informal language. On the other hand, textbook discourse is unique to science and systematic. In line with this, Wallace (2004) conducted ABI research in a seventh grade biology topic, diffusion and osmosis. Students were selected based on some criteria which are their willingness, balanced gender, range of science achievements obtained from the scores before and after the ABI instruction tests. Based on these criteria, 6 students' data were analyzed. Data sources were interviews, ABI reports and letter writing activities. Results of the study showed that three students were labelled as first-hand observers which mean these students mostly focused on hands on activities as their source of knowledge. These students also thought that writing in ABI is important. Moreover, these students thought that when they perform something, they learn permanently. Even though first-hand observers were low-achievers, they actively participated in the ABI process. Moreover, they were able to distinguish claim and evidence. However, these students might have cognitive dissonance when their hands on activity results differ from textbook (authority) results. These students might have positivist epistemology and they might think that results are derived from only hand on activity data. These students have naïve NOS. They could not integrate class explanation and textbook explanation. Their explanations were limited with first-hand data and not supported by textbooks. Likewise, they did not use theories. These students either focused on their observations or provided scientifically incorrect explanation. One of the students was labelled as reader. This student focused on textbook explanation and class discussion as evidence. The student reported facts, concepts, principles, and explanations and did not focus on ABI writing and hands on activities. This student also did not enjoy in participating ABI activity. Student reported that she did not use first-hand data because first-hand data did not explain the process and reasons. She thought that first-hand data is limited with observation. She thought that there is evidence about what happens, but there is no evidence for explanation of the process in observations. This student's personal epistemology shows that scientific events are not explained by solely observation. While student report theoretical information obtained from textbook, student did not use evidence from observations. This student might see that laboratory is verification of knowledge. Moreover, this student might think that textbook knowledge is correct knowledge. Furthermore, student supported this idea by saying that teachers look for this content knowledge in exams. Two of the students who were high achievers were labelled as integrators. These students were able to negotiate different knowledge obtained from discussions, hands on activities and textbooks. These students used numerous sources to get knowledge. Their personal epistemology showed they consider observations and observation coordination with authority (textbook). The study showed that students use sources which are consistent with their epistemology. Positivist students use first-hand data and group idea. Their cognitive mechanisms are "observe, hear, peer opinion, and explaining". On the other hand, reader rejects first-hand observation and thinks that observational evidence is not sufficient for theoretical explanation. Her cognitive mechanisms are synthesis and reflection. Integrators use different type of cognitive mechanisms and they have no cognitive dissonance to integrate textbook knowledge and first-hand data. In this study, both low achievers and high achievers engaged in ABI processes and they improved themselves. However, high achievers negotiated different sources. Therefore, negotiation of different sources (e.g. textbook and observation) should be explicitly taught to students. Reading sources can be much more in ABI studies. Teachers should emphasize that discourses have different nature (group discussion on investigation and class discussion). Grading should be based on both content knowledge and scientific thinking (e.g. distinction of claim and evidence) that results in balanced scientific growth.

In another study, Kim and Song (2006) examined peer argumentation of 8 middle school students. One of the aims of the study was to reveal what type of data students use in argumentation. In data analysis, researchers found codes to answer research questions. Regarding the source of evidence, researchers used two codes which were personal evidence and authoritarian evidence. Findings showed that students mainly used personal evidence including personal experience and knowledge. Evidence was used implicitly in general. Authoritative evidence was based on literature and teacher in general and students used authoritative evidence in order to hide their ignorance. Likewise, students used authoritative evidence to give example when hiding ignorance. Other findings of this study are presented in students' engagement in argumentation process part.

Walsh and McGowan (2017) conducted a study in which scientists provided feedback to students' artefacts about climate change aiming to develop complexity of students' arguments. In this process, students prepared their work, and they revised them based on experts' feedbacks. First, students' knowledge and experience were revealed in climate change. Second, students looked for news and counterarguments in media about climate change. Third, students prepared infographic (i.e., data graphics, poster) about the impact of climate change on species based on their data collection. Last, students showed their infographics and experts provided feedbacks about infographics at least one time in two weeks. A total of 48 students and 14 scientists from different disciplines related with climate change participated in study. Findings showed that students did not use their own data; however, they used data sets, so scientists did not use specific feedback on this practice. Researchers advised that students should use both their own data and other data sets to obtain evidence in their argumentation studies although students of this study only used data sets in this study.

Another study examining data that students used in argumentation was conducted by Kind et al. (2011). Researchers claimed that students usually did not generate evidence and they got the evidence from computer, videos, and second hand data in previous research. If argumentation is integrated in laboratory; alternative hypothesis, data analysis, and evaluation of evidence will be examined. This is superiority of argumentation on other methods. Without argumentation, students follow cookbook lab instructions which do not include discussion and reflection. Argumentation is in line with open ended lab tasks. However, previous research showed that even tasks are argumentation-based; students' argumentation level was low. The more time allocated to argumentation is, the better quality of arguments students has. According to researchers, Students' generation of evidence increases quality of arguments comparing with other evidence which was already existent. Students' hypothesis and interpretation of evidence increases because students owned the evidence which they generated. But, this idea was not so much tested, so one of the aims of this study was to understand whether already existent evidence or student made evidence is better to reveal high quality of arguments. Previous studies showed that tasks were script based, and their solutions are certain. Researchers explained that unproblematic data and clear response cause limited argumentation in lab. Thus, this study prepared three tasks to stimulate argumentation. First task is "complex data" that do not include clear answer. This task was based on heat transfer through radiation. Second task is "conflicting hypothesis". Conflicting hypothesis was provided to make students have different ideas. Before argumentation, students discussed the ideas as true and false, and then they tested the ideas through experimenting. Dissolution of salt in water was task, and related misconceptions were given as ideas. Third task was the extension of first task, "post-investigation discussion", but in this task students did not conduct experiment; on the other hand; fictional data were provided to students. Fictional students used different methods. Students were expected to discuss on method and data. The difference between first and third task was used to understand whether student generated evidence or already available evidence is better to produce high quality argument. Moreover, researchers focused how these three tasks affected students' argumentations. These tasks can be thought as strategies to increase students' argumentation levels because study focused on how students' arguments differed in different tasks. Participants were three classes of students. Four groups, each of three students were selected for analysis. Two groups performed only last task as control group to compare efficacy of student generated data and second-hand data. Data were collected in two weeks. Investigations were video recorded, and written reports were used in analysis. Tasks were introduced in first 10 minutes. There was no explicit argumentation instruction in this study. Students were requested to make evidence based decision and conduct experiment about tasks. How tasks were solved by students was analyzed based on students' orientations. First one was "experimenting" where students focus on data collection and material. Second one was interpretation of the conclusion obtained from experiment that "hypothesizing". Last one was the "co-ordination and evaluation" that was the combination of first two orientations. Findings showed use of type of argumentation (i.e., students' orientations) changed based on groups and tasks. Student made more explanation in written task (task 3) comparing with the task they experimented. They produced more arguments in task 2 that is conflicting hypothesis. Moreover, students reached high quality argument in conflicting hypothesis task. They used data and produce rebuttals in conflicting hypothesis because they needed to solve confliction. Level of arguments was mostly level-2 that is claim supported by evidence. The fewest number of arguments was in post-investigation discussion. This might be caused by limited time allocated to third task. Although students allocated most of their time to the experimenting as data collection, their type of arguments was rarely experimenting. In these limited arguments about experimenting; students focused recording data and its relation to claim. Very few arguments were about the examination of validity of methods. Coordination and evaluation orientation (combination of experimenting and hypothesizing) included most arguments. In these arguments, students evaluated the evidence obtained from experiments. High quality arguments were proposed in hypothesizing because students discuss validity of alternative claims. However, these kinds of arguments were not common. Students did not produce argument when they collect data. Arguments were proposed in explanation of phenomena or evaluation of the evidence obtained from experiments. Over simplification of methodology in tasks caused students not to investigate method. Moreover, students did not seek quality of evidence. This caused students thought that science is simple. Even though tasks were prepared to make argumentation, students spent most of their time in data collection rather than argumentation. Conflicting hypothesis task yielded better results than complex data in terms of number of argumentation although it depends on the groups. Post-investigation discussion task findings were well, but students' perception about right answer and right method decreased the quality of arguments. Comparison of task 1 and task 3 (student generated evidence vs. available evidence) did not result in favour of task 1 surprisingly because students produced more arguments in task 3. Students' wrong perception to science might decrease the quality and number of students' arguments in both tasks because they thought that there is a right answer and right method, so there is no need to make deep argumentation in solution of tasks. If their epistemic understandings increase, their engagement in argumentation based laboratory may be enhanced. They should give up the idea of truth and support the idea of uncertainty in science. If students' epistemology does not change, their experimental process becomes mechanistic. In this study, students could not pass from one orientation to another (e.g., experimenting). If they conduct experiment, their orientations are experimenting. If they do not conduct experiment, their orientations are hypothesizing. Researchers claim that students need to make argumentation more in experimentation on the accuracy of method and data. If data is complex, they could not understand design and methods. On the other hand, if data is simple, students tend not to make argument. This is the dualism that needs attention. Finally, researchers propose that if epistemological criteria are taught to students before investigation, students can benefit more from argumentation based laboratory activities.

### 2.3.3. Summary of Argumentation Studies on Students Epistemological Beliefs.

In conclusion, although there is a theoretical consistency between argumentation and epistemological belief, researchers who conducted ABI approach did not focus on directly epistemological beliefs apart from Tucel (2016). In this study, researcher found that experiment group outperformed control group in justification and development aspects of epistemological beliefs. Researchers adopting socio-scientific orientation, on the other hand, focused on middle school students (Khishfe 2014), high school students (Khishfe et al., 2017), and primary teachers' (McDonald, 2010) NOS views and argumentation. While Khishfe (2014) found that explicit NOS instruction and argumentation treatment supports middle school students' NOS views, McDonald’s (2010) study and Khishfe et al.’s (2017) study yielded blurred relationship between argumentation and NOS. Moreover, although it is not directly related with NOS or epistemological beliefs, Mendonça and Justi (2014) focused on the relationship between scientific models and argumentations and examine students’ arguments. Researchers also focused on type of data that students use in argumentation process and students’ choice of data is linked with their epistemological beliefs. The reviewed articles suggested that students mainly used one type of data. For example; students mainly used either first hand data or second hand data in Wallace (2004) study. Likewise, students used usually personal data in Kim and Song (2006) study. On the other hand, students used second hand data in Kind et al. (2011), and Walsh and McGowan (2017) research. Students are expected to use both first hand data and second hand data when they construct argument and they are expected to make negotiation between first hand and second hand data (Wallace, 2004). Therefore, using only one type of data was linked students’ naïve epistemological beliefs in argumentation.

### 2.3.4. Reflections from Literature to Current Study regarding Epistemological Beliefs and Argumentation.

In this sub-section, how previous literature contributed on current research about argumentation and epistemological beliefs is explained. According to McDonald and McRobbie (2012), students’ epistemological beliefs should not be absolutist because absolutist people think that there is one truth and this truth does not change. When people think that the truth is certain and does not change, they tend not to discuss. Similarly, McDonald and McRobbie (2012) thought that people having multiplist beliefs do not engage in argumentation process because they do not know the value of evidence in argumentation. On the other hand, evaluativist people evaluate claim and evidence. They know that knowledge is evidence based theory. In line with this, when students were trained prior to the study, their beliefs were aimed to be transformed to the evaluativist epistemology in this study. Therefore, some parts of the presentation included information about scientists, how science is done, and characteristics of scientific knowledge (e.g., empirical based). By this way, it is thought that students' epistemological beliefs might shift to evaluativist one. Similarly, Martin and Hand (2009) claimed that when students conduct experiments and discuss on the ideas, they learn how scientific knowledge is constructed. Therefore, when students engaged in argumentation process, their epistemological beliefs improve. In line with this idea, students were told that they work like scientists in this argumentation study when the study was introduced them at the beginning. Role of experimentation and discussions were frequently emphasized in lesson plans too. Another contribution of the literature is about what this study should consider. The alternatives were the epistemological beliefs and nature of science. According to Khishfe (2012), nature of science (NOS) comprehends only one part of epistemological beliefs. NOS deal with nature and construction of knowledge. NOS answers what constitutes knowledge, how knowledge is known, whether knowledge can be known with certainty; however, NOS ignores how individual come to know, the beliefs they hold about knowing, and how epistemic belief influence thinking and reasoning. Therefore, this study specifically focused on students' epistemological beliefs instead of nature of science. Literature review also informed this study about how students' epistemological beliefs are measured. Accordingly, Epistemological Belief Questionnaire (Conley et al., 2004) included all dimensions of epistemological beliefs which are source, certainty, justification and development. This scale was used in previous ABI research examining students' epistemological beliefs development (e.g., Tucel, 2016). Hence, the same questionnaire was used in current study to understand changes of students' epistemological beliefs. Regarding students' epistemological beliefs, previous studies also affected current study in terms of characteristics of data. For example, Kind et al. (2011) reported that students do not understand method and design if data is complex. Likewise, if data is simple, students do not have to engage in argumentation. Therefore, in this study, data should be neither complex, nor simple to engage students in argumentation. Accordingly, new experiments were not produced because they could be either simple or complex. On the other hand, experiments from curriculum consistent with students' grade level were used in this study. These experiments were recipe type or cook-book experiments, but they were restructured consistent with nature of ABI in this study. By this way, simplicity and complexity of data problem was aimed to be solved. Similarly, Walsh and McGowan (2017) suggested that students should use both first hand data (i.e., they data they produced) and second hand data (i.e. the data they obtained from somewhere else). In line with this, current study was prepared such a way that students could use both first hand data and second hand data. Accordingly, students used first hand data and conducted experiments in first, second, fifth and last weeks of the study and second hand data in third and fourth weeks of the study to provide evidence to support their arguments. Kind et al. (2011) also reported that even though researchers design the tasks for students' engagement in argumentation, students focus on experimentation and they do not engage in argumentation. Therefore, time limit was used in this study regarding data collection procedure. Accordingly, students collected their data (e.g., experimentation) in a lesson hour (40 minutes) for each week. Likewise, whole class discussion lasted a lesson hour for each week. By this way, time effect was removed from the study and comparing students' performance regarding argumentation schemes and their engagement in argumentation process in different weeks became possible.

### 2.4. Argumentation and Science Process Skills

This part specifically focuses on firstly, the theoretical link between argumentation and science process skills and secondly, the practical research on argumentation examining its impact on science process skills. This part concludes with a summary of the related research findings.

One of the general aims of schooling is to make students think (Padilla, 1990). Due to fact that science emphasizes making hypothesis, manipulating the physical work and reasoning from data, students' science related skills improve when they engage in science practice. According to Padilla (1990), science related skills show themselves in the terms which are scientific method, scientific thinking, and critical thinking. However, the term science process skills became popular after Sanderson, and Kratochvil (1971) prepared Science--A Process Approach (SAPA). According to SAPA, science process skills are the results of the classification of the science's intellectual tools. These process skills are basic and integrated process skills. Basic process skills are observing, classifying, using numbers, measuring, using space/time relationships, communicating, predicting, inferring. Sanderson and Kratochvil (1971) reported that basic process skills are taught in primary grades. On the other hand, integrated process skills are defining operationally, formulating hypotheses, interpreting data, controlling variables, experimenting. Integrated process skills are taught in intermediate grades (Sanderson & Kratochvil, 1971). This study specifically focuses on integrated process skills as science process skills because this study is conducted with students enrolled in 6$^{th}$ grade level which is one of the intermediate grades.

### 2.4.1. Theoretical Link between Argumentation and Science Process Skills.

Science process skills encompass rational and logical skills used in science. The use of these skills let students to construct scientific knowledge and produce solutions to given problems (Burn et al., 1985). Accordingly, scientific inquiry is thought to be closely related to science process skills because when students make inquiry, they act like scientists and follow scientists' thinking patterns which involve the utilization of various science process skills. In other words, it is expected that students use their science process skills during the inquiry process (Bunterm et al., 2014; Walker & Sampson, 2013). As a result, argument-based inquiry, which is the focus of the current study, is likely to improve students' science process skills because the ABI requires students to formulate their own research questions, propose hypotheses, identify variables design experiments, and discuss the findings These processes have potential to directly contribute to the development of three of the five process skills investigated in this study, namely identifying variables, formulating hypothesis, and designing experiment. Throughout argumentation process, students are also expected to clearly communicate with others. They are supposed to clearly explain their ideas, claims, and justifications. Accordingly, the variables that they investigated should be clear in their minds and they should be able to provide operational definitions. Therefore, the ABI is also likely to improve students' ability to operationally define. In a similar main, students prepare tables when write their reports in ABI. For example, they prepared tables showing information how temperature of water changes over time in matter and heat unit. Such kind of tables might inform students to understand relationships between variables (e.g. temperature vs. time) and students can transform these tables into graphs to better interpret relations between variables. Due to fact that students are familiar to interpreting graph from earlier units of same grade level (e.g., Force and Motion), they can use the data obtained from tables they prepared and draw graphs representing relationships. After drawing these graphs, students can interpret them and they can improve their interpreting graph skill. In conclusion, ABI approach might theoretically increase all integrated process skills.

#### 2.4.2. Argumentation Studies on Students Science Process Skills.

Development of science process skills were usually investigated in inquiry studies (e.g., Bunterm et al.; 2014). However, the number of such studies focusing on argumentation is quite limited. In this part, firstly inquiry studies and their effect on science process skills are presented. Then, the argumentation studies examining the development of science process skills are reported. Moreover, science process skills are closely related with science practices. According to Next Generation Science Standards [NGSS] (2013), the word “practice” is used instead of “skills” to emphasize that scientists engage in scientific investigation. Moreover, practice is not limited with skill, and also includes knowledge specific to practice. Some of the practices in science are planning and carrying out investigations, analysing and interpreting data, obtaining, evaluating and communicating information. Due to fact that these science practices are closely related with science process skills, argumentation studies investigating participants’ science practices are also included in this part of the review.

Regarding the effect of inquiry approach on science process skills; Germann (1989) examined the effect of directed-inquiry treatment on high school biology students’ science process skills. This directed inquiry approach included six efficient instructional and learning methodologies which are learning cycle, focusing technic (i.e., Teacher focuses on student attention to SPS, students use skills alone, responses are discussed, teacher model existing and new skills.), Vee diagram, advance organizers, concept maps, and writing process including prewriting, drafting, revising, editing, publishing, and evaluation. The study consisted of 20 lab exercise and lasted seven months. On the other hand, control group performed the same lab activities, but directed-inquiry approach was not used in this group. Students' science process skills were measured by the use of Test of Integrated Process Skills and Process of Biological Inquiry Test. ANCOVA was used to examine the effect of directed-inquiry approach on science process skills while controlling for students' content knowledge scores. Findings suggested that there was no significant difference between groups in sub-sequent analysis of ANCOVA. Descriptive statistics showed that experimental group performed lower than control group in all tests Researcher also conducted regression analysis to reveal possible aptitude treatment interaction and results showed that there was an interaction between students' cognitive level and type of treatment they took. More specifically, directed-inquiry approach was found to be more effective for students classified as concrete operational learners while less guided approach (i.e. control group) was more effective for high achievers or students in formal operational thinkers. Therefore, directed inquiry, its guidance and slow pace teaching might preclude high achievers' cognitive strategies and further learning. Researcher added that the best instructional methodology for average person may not be best for all students. Researcher also claimed that assigning more students in formal operational stage to the control group might cause the performance difference in favour of control group.

Likewise, Bunterm et al. (2014) examined the impact of different types of inquiry, which are structured and guided inquiry, on 7$^{th}$ and 10$^{th}$ grade level students' content knowledge, science process skills, scientific attitudes and self-perceived stress. The study lasted 14-15 hours and 239 students participated in the study. Researchers claimed that although inquiry and traditional teaching instructions were compared in previous research, there was no study examining the impact of different types of inquiry (i.e., structured vs. guided inquiry) on students' achievements and other related variables in literature. In structured inquiry, researchers provided research questions and instruction regarding how students complete the experiments. On the other hand, students designed and tested their own experiments, and reached conclusion in guided inquiry, but research question was given them. In other word, the difference between structured inquiry and guided inquiry was designing and testing part of experiment. While guided inquiry let students be free in designing and testing experiment, structured inquiry provided the instruction that students were expected to follow in their designing and testing part of experiment. Researchers used 5E model to create inquiry classes. Science process skill test included 45 multiple choice items and 13 science process skills. The test was independent from the topic covered in the study (i.e., work, energy and motion). Data was analyzed by use of multiple analysis of variance and content knowledge, science process skills, science attitudes and stress were used as dependent variable. Findings showed that although students' science process skills improved significantly in structured inquiry, it improved more in guided inquiry. The same trend was also observed for students' content knowledge, however, difference in effect size was smaller between guided inquiry and structured inquiry regarding content knowledge. Although students' attitudes towards science improved, there was no significant difference between guided inquiry and structured inquiry regarding attitudes. Likewise, students' stress decreased equally for different inquiry types from pre-test to post-test.

In Turkish context, Koksal and Berberoglu (2014) examined the impact of guided inquiry approach on $6^{th}$ grade level students' content knowledge, science process skills and attitude towards science. There were 304 students who participated in the study in total. Experimental group included 162 students and control group had 142 students in this quasi-experimental study. The study was conducted in Reproduction, Development, and Growth in Living Things topic. The study lasted 22 hours in the study. Teacher was the key factor and guided all process in the experimental group taught by guided inquiry. During the guided inquiry process, teacher did not give the answers and present steps of inquiry. On the other hand, teacher provided clues for students to follow. Thus, a semi-structured approach was used. Researchers claimed whether guided inquiry approach is efficient for science process skills is not well known. Control group teachers followed teacher guide book suggested by Ministry of National Education. Activities that teachers conducted in control group were based on verification of scientific knowledge, and not related with inquiry even though these activities were experiments. Science process skills was measured by use of SPS test and this test included observing, classifying, proposing hypothesis, identifying variables, proposing data, and producing models. This test included 16 questions which are multiple choice items, open-ended items, matching and hotspot items. Content of the science process skills test was not specific to Living Things topic. Repeated analysis of variance was used to answer the research questions. Findings suggested experiment group performed better than control group in science process skills, science achievement and attitudes towards science. Although effect size was small for science achievement, it was medium for science process skills and attitude.

In another study conducted with middle school students, Durmaz and Mutlu (2017) focused on the effect of instructional intervention on 43 seventh grade students' science process skills. Data collection process lasted 18 weeks. Students' science process skills were measured by use of Science Process Skills test including 40 questions and 13 aspects of science process skills which are observing, classifying, drawing conclusions, predicting, measuring, recording data, space/time relation, operational definition, formulating hypothesis, designing investigations (experimenting), controlling and manipulating variables, analysing data and constructing model. Reliability coefficient of the scale was found as 0.81. Experiment group intervention included a treatment that emphasizes science process skills whilst control group did not emphasize SPS. Both experiment and control group conducted similar activities including carrying out experiments, observations, measurement, recording data, and research questions as suggested in curricular sources (i.e., textbook). Moreover, experiment group was scaffolded with worksheets and experiment reports and teacher provided feedback on students' work. Furthermore, experiment group's activities started with closed experiments conducted by the teacher and continued with partial open ended experiments that students and teacher share the responsibilities in conducting experiments, and finished with open ended experiments that students control the experimenting process. Paired sample t-test results showed that there was a significant improvement in the experiment group. On the other hand, there was no statistically significant change in control group' science process skills from pre-test to post-test. Moreover, descriptive analysis showed that experiment group developed most of the science process skills throughout the study. Developed SPS aspects were identifying research question, formulating hypothesis, identifying and controlling variables, determining materials, designing experiments, tabulating data, and drawing conclusions. Although students developed these SPS aspects, researchers noticed many problems about SPS. For example; students focused on the conclusion of research problem rather than the reason, students confused dependent, independent and control variables. Likewise, students could not link the variables when they formulate hypothesis. They used more than one independent variable. They also had difficulty in drawing graphs. Students also had difficulty in connecting their conclusion with hypothesis.

In his master thesis, Yıldırım (2012) examined the effectiveness of a guided inquiry treatment on 55 8$^{th}$ grade level students' science process skills and content knowledge in physics topics including buoyancy and pressure. In this quasi experimental research, there were three groups and two of them were experiment group and one of them was control group. While experiment group included 39 students, control group included 17 students. Although, numbers of participants in experiment group and control group were highly different, researcher reported that test results met all the assumptions (e.g., equality of variance) to conduct inferential statistics. Students' science process skills were measured by use of Science Process Skill test created by Burns et al. (1985). Although experiment groups outperformed control group in science achievement, average gain score of experiment group was not different from control group regarding science process skills. Likewise, researcher found no significant main effect of guided inquiry treatment on students' science process skills. Researcher reported that short period of treatment (5 weeks) might cause no difference between experiment and control group in terms of science process skills. Moreover, long sentences exist in science process skill test might negatively affect improvement of science process skills. Students might not understand the long sentences and they could not perform their science process skills.

Inquiry studies were not only held with middle school and high school students, but also held with teachers. For example; Akben (2015) conducted a case study to understand whether 30 pre-service primary school teachers design different type of inquiry based experiments and use corresponding science process skills. Researcher claimed that activities or experiments suggested in curricular sources (i.e., textbooks) focus on limited number of science process skills. Moreover, these experiments given in textbooks are about verification of facts. However, these experiments should be prepared in different levels of inquiry (e.g. guided inquiry). In Akben's study, participants who are pre-service primary school teachers were not familiar with science process skills, inquiry and its types prior to the study. After theoretical knowledge was given to participants, researcher selected five different experiments from fourth and fifth grade level textbooks. Pre-service teachers were assigned to transform given textbook experiments to the experiments including different types of inquiry. Moreover, they were expected to show to what extent their transformed experiments match with science process skills. After pre-service teachers planned their experiments, they conducted them in class and their peers provided feedbacks. Based on the feedbacks on presentation of their experiments, pre-service teachers made adjustments on their experiments. Findings suggested that pre-service teachers prepared open inquiry experiments related to basic topics about which they were knowledgeable whilst they prepared closed inquiry experiments in difficult topics. To show the link between inquiry implementation and emphasized process skills, researcher selected three different experiments. Topics of the experiments were "change of mass in response to heating", "propagation of light", and "separation of mixtures". Pre-service teachers' explanations suggested that they point only few number of SPS (4-6 science process skills) when they prepare experiment to verify facts. When pre-service teachers prepared closed inquiry, they pointed out 5-8 process skills. Likewise, guided inquiry experiments corresponded to 9-10 science process skills and lastly open inquiry experiments matched with 10-12 science process skills. In conclusion, the findings suggested that pre-service primary school teachers are capable of conducting different type of inquiry experiments and they could use science process skills in their different type of inquiry experiments. Likewise, pre-service teachers were aware that when inquiry level increases from closed inquiry to open; more science process skills are emphasized in science practice.

Even though it is not as popular as inquiry studies, argumentation studies also examined students’ science process skills and science practices. For example, in an argument-based inquiry study, Aslan (2016) examined the impact of argumentation based inquiry approach on 53 pre-service teachers' science process skills and their attitudes towards laboratory courses. In this study, there was no control group and both groups were experiment group: Groups were divided as students having higher scores and students having lower scores based on their science process skills' pre-test results. This assisted researcher to understand how argumentation based treatment influences two groups of pre-service teachers categorized based on their prior science process skills. The treatment included 9 experiments in chemical kinetics, chemical equations, acids and bases, gases, heat, and electrolysis topics. Findings suggested that there was a significant improvement in both groups concerning overall science process skills. Moreover, students having lower science process skills at the beginning of the study developed their overall scores more than students having higher science process skills. Likewise, both groups’ science process skills sub-components’ scores (e.g., operational definition) were similar at the end of the study. In conclusion, students having lower scores at the beginning developed themselves more than others. Regarding attitudes towards laboratory courses, the improvement was statistically significant for both groups. Interview data also suggested that 14 pre-service teachers over 19 thought argumentation based treatment was beneficial because it brings about academic achievement (e.g., meaningful learning) and personal development (e.g., development of thinking skills).

In another argumentation study, Walsh and McGowan (2017) aimed to develop high school biology students’ science practices. In this study, scientists were invited to the program and scientists provided feedback to the students. In this process, students prepared their work on climate change and they revised these works based on scientists (experts) feedbacks. Researchers analyzed student work, scientists' feedbacks and transcripts of student-scientist interactions. Science practices included eight themes which are i. asking question and defining problem, ii. developing and using models, iii. designing and carrying investigation, iv. analysing and interpreting data, v. mathematical thinking and calculation, vi. Deep explanation and proposing solution, vii. proposing arguments based on evidence, and viii. obtaining evaluating and communicating knowledge. Regarding asking question practice, researchers reported that scientists asked questions to assist students. After scientists asked questions to students, students used more evidence and their arguments became stronger. After that, scientists checked whether students' evidence fit with arguments. Scientists also connected their work with students' works and asked further questions about students' work. Then, scientists' questions let students to use more evidence. Scientists also asked prompting questions to the students. When students answered prompting questions, they got new evidence. Scientists' questions also let students to think on multiple perspectives. When students thought on multiple perspectives, they used more counterarguments.

Results about analysing, interpreting data practices and mathematical calculations practices showed that mathematics and data analysis are important because students' arguments should be supported by evidence derived from data analysis. Scientists criticized the data students used because they wanted students to focus on data and its analysis. For example; scientists emphasized deficient analysis. Regarding scientific explanations in argumentation, it was reported that scientists looked for integrity and accuracy of explanations. In this way, students' explanations were clarified and became more academic. Students' revised work included new evidence, and students corrected their mistakes and missing points. Feedbacks also pointed out inconsistencies between evidence and explanations, students corrected these inconsistencies. For the models use in argumentation; scientists supported students' use of models because these models support their arguments (i.e., future maps). Scientists also suggested students to add models which can be helpful in their arguments. Designing and carrying out investigation practice results showed that students did not use their own data; however, they used data sets, so scientists did not use specific feedback on this practice. Finally, obtaining and communicating multimodal arguments is another practice. Researchers reported that students are expected to use multimodal data including visuals, graphs, and texts. Accordingly, scientists suggested students to prepare posters composed of multimodal data such as text and visuals. Scientists also advised students that texts and visuals used in posters should be consistent with each other. Moreover, scientists asked students whether visuals are appropriate for their arguments. Scientists added that some visuals are used to support readability of poster rather than being used as evidence. Moreover, scientists found some evidence for students to support their argument, and they explained students how they found this evidence in feedbacks. Researchers concluded that students focused on one thing in their first work, but they looked from multiple perspectives after they got feedback from scientists. Hence, they considered more evidence. In other word, holistic thinking that they developed produced stronger arguments that consistent with NGSS (2013) scientific practices.

### 2.4.3. Summary of Argumentation Studies on Students Science Process Skills.

In conclusion, previous studies showed that most of the treatments carried out to improve science process skills are related with different types of inquiry (e.g., guided inquiry). Findings of these inquiry studies are different from each other. Although some of them found no difference between inquiry group and control group (German, 1989; Yıldırım, 2012), some other researchers found that inquiry groups performed better in science process skills at the end of the treatments (Akben, 2015; Bunterm et al., 2014; Durmaz & Mutlu, 2017; Koksal & Berberoglu, 2014). Moreover, research showed that when type of inquiry transform from closed to open inquiry participants develop their science process skills more (Akben, 2015). Moreover, Aslan’s (2016) study showed that argument-based inquiry approach is beneficial to improve pre-service teachers’ science process skills significantly. Furthermore, the researcher suggested that pre-service teachers having lower scores improved their science process skills more than their counterparts. Due to fact that science practice includes science process skills, one more study about argumentation and science practice was reviewed. The finding of the study showed that while constructing arguments, students developed their science practices (e.g., analysing, interpreting data and mathematical calculations) with the help of scientists' feedbacks.

#### 2.4.4. Reflections from literature to Current Study regarding Science Process Skills and Argumentation.

Next, how previous literature contributed on current research about argumentation and science process skills is explained. Sanderson and Kratochvil (1971) reported that basic process skills are taught in primary level and integrated process skills are taught in middle school level. Due to fact that this study was conducted with middle school level students, students' integrated process skills were considered in this argumentation study. Similarly, literature review also assisted to determine which instrument was used to measure students' science process skills in this study. Theoretically, ABI is consistent with five aspects of integrated process skills (e.g. formulating hypothesis). Therefore, a scale was needed that measure all these five science process skills. In line with this, science process skills test developed by Burn et al. (1985) assessing all integrated process skills were used in this study. Moreover, previous study on science process skills and different inquiry levels also shaped design of the current study. Accordingly, Akben (2015) suggested that the more students are free in inquiry, the more science process skills they use. Therefore, guided inquiry and open inquiry should be used in inquiry studies rather than cookbook experiments and closed inquiry. In line with this, when the lessons were designed, mainly guided inquiry was adopted. Guided inquiry was also suggested by Ministry of National Education [MONE] (2018) teaching program for 6$^{th}$ grade level. Accordingly, students were active learners, they were generally free, they had rights to ask their own research questions, and conduct their own experiments in this study. By this way, it is thought that number of science process skills that students use and develop would increase. Similarly, cook book experiments and closed inquiry was not used in this study. Likewise, while lessons were prepared, integrated process skills defined by Sanderson and Kratochvil (1971) were considered. Accordingly, there are five process skills which are defining operationally, formulating hypothesis, interpreting data, controlling variables and experimenting. Students were suggested to explain their hypothesis and variables in all weeks that they conduct their experiments. By this way, students were expected to increase their formulating hypothesis, controlling variables, and experimenting process skills. Likewise, some contexts were provided to students that they can produce or obtain data in all weeks. By this way, students found chance to interpret data. For example, they prepared tables to interpret what their data means. Likewise, students presented their experiments to other groups and discussed with each other. When students explain their research, teacher and other students asked students to define the variables that they used in their experiments. By this way, it is possible that students improve their defining operationally process skills. The next section of the literature review is about the argumentation studies focusing on argumentation schemes.

### 2.5. Argumentation Schemes

Sampson and Clark (2006) examined analytical methods used in argumentation studies that aim to understand students' argument quality. These analytical methods are Toulmin Argument Pattern (TAP), Zohar and Nemet's (2002) content of justification in an argument, Kelly and Takao's (2002) Epistemic Levels of propositions in an argument, Lawson's (2003) hypothetical deductive validity of an argument and Sandoval's (2003) conceptual and epistemological aspects. According to Sampson and Clark (2006), on the other hand, students should learn five epistemic criteria, which are not emphasised by aforementioned five analytical methods, to produce high quality argument. These five criteria include examination of (a) the nature and quality of the knowledge claim, (b) how the claim is justified, (c) if a claim accounts for all available evidence, (d) examination how the argument attempts to discount alternatives, and (e) examination how epistemological references are used to coordinate claims and evidence. In line with the idea, Duschl (2007) proposed that Walton's (1996) argumentation schemes consider epistemic criteria while assessing students' arguments. Therefore, it is expected that using argumentation schemes provide more valid and reliable results regarding quality of students' arguments. In line with this, argumentation studies using argumentation schemes which are sign of high quality arguments are reviewed in this part. In this point, defining Walton's argumentation schemes can be meaningful.

According to Walton et al., (2008), argumentation schemes are forms of argument. These schemes show structures of common types of arguments that people use in their everyday discussions such as legal argumentation happening in courts and scientific argumentation. There are 25 argumentation schemes that Walton (1996) defined. Some of these argumentation schemes are argument from sign, argument from expert opinion, argument from cause to effect and so on. As arguments can be deductive and inductive, they can also be presumptive. Argumentation schemes include presumptive reasoning. They are not strong arguments because of lack of evidence. Although these arguments having presumptive reasoning are not so strong, they can support the conclusion by providing evidence. In other word, these schemes carry weight, and can affect the discussion moving forward. By the contributions of argumentation schemes, some tentative accepted conclusions are reached even though all evidence is not known. These argumentation schemes are common in everywhere including scientific investigations. Although these argumentation schemes are necessary for reaching a conclusion, these schemes are still fallible and tentative conclusion can change when new evidence comes. For example, an argumentation scheme may propose a generalization and conclusion is shaped by this generalization. After that, an exceptional case is defined and reached conclusion changes.

### 2.5.1 Studies focusing on Argumentation Schemes.

Although argumentation schemes involve epistemic criteria for high quality arguments, they were not used so much in argumentation studies. In one such study, Duschl (2007) worked with middle school students to improve their argumentation skills through SEPIA (Science Education through Portfolio Instruction and Assessment) program by using argumentation schemes. In SEPIA study, researchers focused on whether middle school students' scientific reasoning improved in terms of ability to reason about explanations, experiments and models. Students were expected to evaluate claims using epistemic criteria; therefore minimum content knowledge was used in this study. Seventeen triads of middle school students participated in one hour long interviews in this study. Students evaluated the scientific posters. Interviews were videotaped and transcribed. At first, both TAP and Walton scheme was aimed to be used as analysis method, but TAP was abandoned. Walton (1996) identified over 20 different argument categories. Nine of these argumentation schemes (sign, commitment, position to know, expert opinion, evidence to hypothesis, correlation to cause, cause to effect, consequences, and analogy) are appropriate for middle schools. Revealing argument categories assisted researcher to understand how student use evidence. In analysis, researchers could not negotiate on the categorization of students' arguments and they decreased these categories from 9 to four which are: request for information (sign and position to know schemes), expert opinion (expert opinion schemes), inference (correlation to cause, cause to effect, consequence, evidence to hypothesis), analogy (analogy schemes). Inter-rater reliability of a transcript was calculated as 90%. SEPIA group members produced more dialogical argumentation schemes than non-SEPIA group members. Likewise, SEPIA group produced more number of arguments. According to descriptive results, SEPIA group members produced more inference and request for information categories. Comparison of nine argumentation schemes for both groups showed that groups had similar ranking in terms of argumentation schemes where Spearman Rank Correlation Coefficient was found as 0.95. This study showed that middle school students are ready to do presumptive reasoning. Walton argumentation schemes showed that individuals' performance in argumentation discourse is more than what TAP suggested. This study also showed that presumptive reasoning analysis is important to understand and evaluate development of students' argumentation strategies.

Duschl (2007) also claimed that science is not purely logical and analytical as emphasized in textbooks; rather, it is dialectical and rhetorical in line with process of scientific research. Hence, dialectical context can be more important to engage students in science classes. Moreover, dialectic argumentation includes presumptive reasoning. Presumptive reasoning is done when there is need for action and evidence is not sufficient. Burden of proof changes throughout the presumptive reasoning. Therefore, Walton's Argumentation Schemes for presumptive reasoning is useful to analyze middle school students' arguments.

Although it is not directly related with scientific arguments, Macagno, Mayweg-Paus and Kuhn (2015) aimed to increase students' arguments by using expert attacks. Sixteen middle school students participated in the study and they discussed on a pedagogical issue as a pair. Then, experts (PhD students) substituted with opposing peer in pair and joined to the electronical discussions as mock students in this study. Students' arguments changed after student and expert interaction. After expert and student interaction, students focused more on values and they revealed more solutions to the problem. Next, experts were removed and students re-discussed on the solution of the problem. Students started to discuss their peer's principles forming the argument rather than directly attacking arguments. Six different argumentation schemes were revealed based on the content of the arguments which are argument from consequences, practical reasoning, argument from values, argument from best explanation, argument from rules, and argument from classification. One of the limitations of this study was the experts who were PhD students, not teachers. Moreover, this study was not focused on science education. On the other hand, this study focused on general pedagogical issue. This short intervention provided evidence that students develop their arguments in response to treatments.

Argumentation schemes studies were not only held with middle school students, but they were also held with pre-service science teachers. In one of these studies, Ozdem et al., (2013) examined 35 pre-service science teachers' argumentation schemes in different tasks in an argumentation based inquiry treatment. Class discussions were video-recorded and transcribed for data analysis and Walton (1996) argumentation schemes were used as analytical framework to understand pre-service teachers' arguments. In the study, firstly, researchers identified pre-service teachers' argumentation schemes and secondly, they compared the frequency of argumentation schemes revealed in experimentation process and the frequency of argumentation schemes revealed in whole class discussions. Researchers found 20 different argumentation schemes and the most observed argumentation scheme was argument from sign in experimentation process. On the other hand, 18 different argumentation schemes were observed in class discussion. Argument from correlation to cause was the most frequently observed argumentation schemes in class discussions. Argument from sign and argument from evidence to hypothesis followed argument from correlation to cause. Researchers proposed that these three argumentation schemes are regarded as scientific argumentation. Regarding argument from sign, its close relationship with inferences, empirical observation and measurement and graph makes argument from sign as scientific argumentation. Likewise, argument from evidence to hypothesis is an example of scientific argumentation because pre-service teachers made some verification and falsifications in their laboratory work. Similarly, pre-service teachers observed the relationships between variables and they produced argument from correlation to cause by using their inductive reasoning in discussions and experimentation. Researchers added that argument from position to know, ethotic argument, the causal slippery slope argument, the precedent slippery slope argument and verbal slippery slope argument were not used by pre-service teachers. Pre-service teachers' participation to argumentation differed in tasks. For example, they engaged more in density task. While argument from correlation to cause was most observed argumentation schemes in three tasks, argument from sign was the most observed scheme in other three tasks. On the other hand, each task included similar number of kinds of argumentation schemes ranging from 12 to 17 argumentation schemes. Researchers also suggested that argument from sign, argument from example, argument from evidence to hypothesis, argument from correlation to cause, argument from cause to effect, and argument from consequences were observed in all tasks and these argumentation schemes can be task independent. Researchers also added that experimentation stage provided participants a context where they can construct their arguments and class discussion provided a context where they could evaluate the arguments. Experimentation stage lasted more than whole class discussion. However, researchers observed that number of arguments that participants produced in class discussion were more than it was expected. Researchers also reported there were differences of the use of argumentation schemes between construction of arguments and their evaluation. Accordingly, pre-service teachers used argument from sign and argument from an established rule more in constructing scientific knowledge. Conversely, they used argument from evidence to hypothesis, argument from expert opinion, argument from consequence, argument from analogy, and argument from bias when they evaluated the arguments.

In another study held with 40 pre-service teachers, Konstantinidou and Macagno (2013) focused on understanding participants' reasoning. Because reasoning is something abstract, to make it concrete, the researchers used argumentation schemes while analysing data. The overarching aim of the study was to remedy participants' background knowledge leading to flaws in their arguments. Researchers focused on warrant component of TAP model to remedy students' arguments. By using argumentation schemes, researchers aimed to clarify premises because premises are abstract and hidden. After categorization of arguments, premises can be better understood. Argumentation schemes include abstract premise and conclusion including warrant. Even though argumentation schemes do not include many components like rebuttal, they focus on the content and logic of the argument. Researchers focused on background knowledge showing clues for participants' reasoning. After understanding pre-service teachers' background knowledge, researchers used this knowledge to reconstruct their arguments. In analysis method, researchers focused on understanding the quality of participants' arguments. While analysing the arguments, researchers reached participants' prior knowledge. Then, they aimed participants to understand their own arguments and improve initial arguments. As human beings, we use some criteria to make reasons. For example; our prior knowledge could be criterion. Likewise, warrant can develop based on laws, statistics, authority, causal relations, and ethical commitment. Arguments were classified depending on its warrants (9 class including cause and analogy). In thermal conductivity activity, wrapping a snowman in a coat, researchers showed a cartoon and provided three alternative ideas. After selecting one of the alternatives, participants choose the most important terms and define them. Different definitions were discussed in class. Individual arguments about cartoon were constructed in 20 minutes whereas formation of group arguments lasted 1 hour about the thermal conductivity activity. Participants' written arguments were analyzed. Findings suggested that %80 of the arguments were cause-effect argumentation schemes and %10 of them were analogical argumentation schemes. The reason why participants mostly explained causal relation was related with participants' education level. Research showed that undergraduate level students mostly use causal arguments based on their prior knowledge even though some conclusion can be wrong. Causal arguments are stronger than analogical arguments. Previous research showed that earlier education level students produce analogical arguments which depends experience more than prior knowledge.

### 2.5.2. Summary of the Argumentation Schemes Studies.

In conclusion, although there are few studies on argumentation schemes held in science classes, these studies inform us about students' use of argumentation schemes in their discussions. More specifically, results suggest that dialectical context can promote use of argumentation scheme in science classes, middle school students are able to do presumptive reasoning and the ones having argumentation treatment are better than their counterparts in using argumentation schemes (Duschl, 2007). Moreover, the most used argumentation schemes were reported as inference (correlation to cause, cause to effect, consequence, and evidence to hypothesis) and request for information (sign and position to know) by Duschl (2007). Similarly, Macagno et al.'s (2015) use of argumentation schemes showed middle school students' arguments develop as response to expert attacks. Regarding use of argumentation schemes in undergraduate level, Ozdem et al. (2013) study which used argument based inquiry treatment showed that pre-service teachers mostly use argument from sign schemes during experiments, and they use argument from correlation to cause schemes in class discussions at most. Researchers thought the most used three schemes which are argument from correlation to cause, argument from sign and argument from evidence to hypothesis are scientific argumentations. Moreover, findings of this study pointed out that argumentation schemes that pre-service teachers used changes depending on the tasks; however, number of total argumentation schemes does not vary among different tasks. Furthermore, pre-service teachers focus on constructing arguments in experimenting and evaluating arguments in class discussions. Researchers reported that pre-service teachers use arguments from sign and established rule when they construct arguments. On the other hand, they use argumentation schemes which are argument from evidence to hypothesis, expert opinion, consequence, analogy, and bias when they evaluate others' arguments. Finally, Konstantinidou and Macagno's (2013) study showed that pre-service teachers mainly use cause-effect argumentation schemes in their arguments. Why they usually use cause-effect schemes were explained in terms of their grade level and prior knowledge.

### 2.5.3. Reflections from Literature to Current Study regarding Argumentation Schemes.

Next, how previous literature contributed on current research about argumentation schemes is explained. Sampson and Clark's (2006) analysis was informative for selection of data analysis regarding argument quality. Accordingly, Sampson and Clark (2006) reported that the studies examining argument quality do not meet the epistemic criteria (e.g. examination how the argument attempts to discount alternatives). By pointing out this analysis, Duschl (2007) reported that Walton's (1996) argumentation schemes meet epistemic criteria and these schemes are suitable to analyze participants' argument quality. In line with this, Walton's (1996) argumentation schemes were used for analysis to understand quality of arguments in this study. Likewise, coding was prepared based on the work of Walton (1996) and Duschl (2007). For example; Walton's (1996) examples were related with daily life and therefore, they were not sufficient to analyze scientific arguments' analysis. In this point, how Duschl (2007) analyzed same codes was examined because Duschl (2007) analyzed middle school students' scientific arguments. By this way, coding procedure became clearer. For example; Walton's explanation about argument from corelation to cause just says that a corelation between two variables is seen like there is causation. On the other hand, Duschl explained same argumentation scheme (i.e.

argument from corelation to cause) in detail and Duschl says that people prefer plausibility to rather than possibility in argument from corelation to cause. In conclusion, using explanations of codes for argumentation schemes from two different sources facilitated the argumentation schemes analysis in this study. Lastly, unit of analysis to analyze participants' argumentation schemes was borrowed from Duschl (2007) and unit of analysis was reasoning sequence in this study in line with Duschl (2007). The next part informs about the studies examining students' engagement in argumentation process.

## 2.6. Argumentation Process

Previous research mainly focused on quality of students' arguments, but how students learn arguing, support and refine their arguments is not clear. Therefore, focusing on process of argumentation is as valuable as studying on quality of arguments (Kim & Song, 2006). In line with this need, science education researchers changed their agenda into argumentation process to reveal nature of argumentation process that results in students' arguments. This part reviews the previous studies investigating students' engagement in argumentation process.

### 2.6.1. Studies focusing on Argumentation Process.

In one of the studies focusing on argumentation process, Cavagnetto et al., (2010) examined fifth grade students' small group interactions in four different units which are terrestrial biomes, aquatic biomes, light and sound using argument-based inquiry approach. Researchers focused only on group talks but not on the whole class discussions. In analysis, students' talk was categorized as on task and of task talk. While on task talk shows students' talking about generating and evaluating arguments, of task talk represents talking which are not related with argumentation. On task talk was also divided in two parts which are Generate talk and Representational talk. Generate talk included students' knowledge construction process and arguments' components were obtained from this talk. On the other hand, Representational talk was related with students' application of Generate talk on writings. Researchers also examined the type of languages to understand functions of students talk. There were seven different language types including instrumental, regulatory, interactional, personal, imaginative, heuristic, and informative. In instrumental language, people satisfy their needs. This language is most used language among people. For example, students may claim that they do not want to study at the end of the lesson. This shows that students satisfy their needs during class. Students control others when they use regulatory language. Interactional language shows students' relationship with each other. If a student talks about herself, this language is personal language. Imaginative language emerges when students creatively use the sounds in their talk. On the other hand, heuristic language includes students' curiosity, exploration and questions. Finally, informative language is used when students provide information to others. Researchers claimed that using heuristic and informative language is hierarchically better than other five language types. Findings of the study showed that students usually used On task talk that shows students talked about generating and evaluating arguments and students focused on knowledge construction process in their small group discussion. Throughout the units, students talk included both generative talk that means they constructed arguments and representational talk meaning they reported their arguments in writing. Regarding components of an argument, fifth grade students usually mentioned claim and data, but they hardly used rebuttals and counterclaims. Rebuttals were in question format and they did not include explanations. Few numbers of rebuttals were resulted with limited number of warrants and backings decreasing quality of arguments. While students' claims and data were used explicitly in talks, other components of argument were used implicitly. Regarding functions of language, %78 of students talk was informative and %12 of talk was heuristic. This shows that students actively participated in the process of generating claims and evidence. Likewise, students seldom used other language types that do not correspond with argumentation in their talk. Researchers added that use of claims and data indicated active knowledge construction by students, however limited number of rebuttals and counterarguments were sign of students' deficiencies on criticizing arguments. Moreover, researchers thought that students were able to transfer their knowledge construction process to different science units (e.g., terrestrial biomes); therefore, argumentation process may not differ across different topics. Finally, researchers claimed that achievement in argumentation studies depend on students' perception of culture. Students' perception should change from doing the lesson to learning through science discourse for more efficient argumentation lessons.

In another study held with fifth grade students, Choi, Hand, and Norton-Meier (2014) examined fifth grade students' participation of an online argumentation treatment using ABI. In this study, there was an asynchronous online discussions forum and students wrote this forum their ideas in plant and health topics during discussions. While students wrote 739 notes in plant topic, they wrote 686 notes in health topic as their arguments. Moreover, ABI was used while students engaged in inquiry based investigations. A total of 129 students participated in the study. Findings suggested that students actively engaged in online discussion. Researchers detected some codes about students' contributions to online argumentation. Accordingly, students "provide simple answer, provide more answer, justify, clarify, accept, challenge, request clarification, challenge test questions, agree with claim, request a detailed claim, disagree with claim, propose a counter claim, challenge question-claim relation, support evidence, provide evidence, challenge evidence, request more evidence, query about backing, challenge reference resource, provide reference sources, request reference sources, challenge test, query about test, provide information about test". Accordingly, more than 85% students actively participated in online discussion. Students' participation in different investigation (human health and plant) was similar. During this process, students focused on the evidence rather than claims. Accordingly, students specifically used evidence in 31 % of all online discussions in plant topic and 48 % of online discussion in human health topic; On the other hand, students focused on claims only in 19 % of online plant discussion and 29 % of online human health discussion. Students proposed some evidence to online discussions to support their claim. After that, many students actively criticized the evidence proposed by their friends. While students criticized the evidence they focus on sufficiency, reliability, validity and accuracy of evidence. On the other hand, students did not so often request more evidence, query backing, provide more evidence and support evidence. When presenting students get critics about their evidence, they both provide simple answer and more evidence. On the other hand, when presenting students defended their claims against responding students' counterarguments, they tried to make justification to validate their arguments. Likewise, if presenting students thought that their arguments were not well understood, they tried to clarify their arguments. Moreover, when responding students attacked to their arguments, presenting students sometimes re-attack (challenge) the responding students' counterarguments. Similarly, when responding students propose counterarguments and attack to presenting students' arguments, presenting students often requested clarification for the responding students' counterargument. In this study, researchers found many codes while students engaged in online argumentation such as "request clarification, challenge test questions, agree with claim". However, students' uses of these codes differed from plant investigation to human health investigation. In plant investigation, students frequently used the following codes in online argumentation: "challenging, providing information, querying test design, procedures, and measurement data". On the other hand, students focused on the challenge reference resources of evidence in human health investigation. Researchers explained these differences based on the differences between investigations. Accordingly, students were expected to generate testable questions in plant investigation, but they were expected to use data from written sources in human health topic. Likewise, research question was supplied to students before argumentation in human health. In line with this, students challenge question-claim relationship in plant investigation, they did not challenge relation of these two argument components in human health investigation because question was already given them in human health topic. In conclusion, it can be said that whether students generate their own question, collect their own data, use already available evidence caused differences on students' engagement in argumentation process.

In a similar vein, Kim and Hand (2015) focused on how elementary science teachers' characteristics and their students' argumentation are related. In argumentation discourse, role of the teacher was to sustain argumentation in class. ABI was selected as strategy to generate arguments in class. Modified Chen's framework (2011) was adopted in the study which included construction (i.e., information seeking, elaboration) and critique (i.e., challenging, rejecting, defending, and supporting) components of oral argumentation. Thirty teachers participated in ABI instruction. Treatment group explicitly got ABI instruction, practiced on it, and collaborated with their colleagues. On the other hand, control group's instruction did not explicitly refer ABI although content of the instruction was related with ABI. After teachers' courses, 6 of them were selected to determine characteristics of teachers using argumentation. Observation protocol was helpful to select participants. Accordingly, treatment group teachers did not get low scores, and control group members did not get high scores. Thus, 3 high score teachers and their classes, and 3 low score teachers and their classes were chosen to compare characteristics of teachers and events in argumentation class. Six hours of videos were analyzed. The grade levels of classes were 3 and 5. Each codes' frequency (e.g., challenging) was summed up, their percentage over total number of codes was obtained for each class and they were averaged in order to compare codes between traditional and argumentation class. Likewise, teachers' movement and interaction were divided into lecture time to get percentage of teacher movement. Four different characteristics revealed after analysis. First one was related with structure of teacher and student argumentation and student-student argumentation. Second is directionality; teacher says students what to do in class and non-directionality, teacher allows students to study alone in activities. Third is movement; when teacher moves in class, s/he interacts with students. Final one is structure of student talk referring argumentation discourse among students. This component focused on critique aspect of argumentation rather than construction. Finding of the study suggested that treatment group teachers used critique component more than control group teachers regarding structure of student and teacher argumentation. Control group mainly reiterated the knowledge by focusing construction component. Control group also did not use evidence in their arguments. Regarding directionality, treatment groups' teachers did not direct their students; they more frequently let students to say their ideas. Moreover, these teachers allowed longer duration for students' speech. On the other hand, control group teachers used more directives. Regarding movement, treatment group teachers moved and interacted with students more than their counterparts. Structure of talk results revealed that students from treatment group teachers used more evidence based reasoning. None of the class addressed challenging dimension. Treatment group teachers' students also used more evidence based support and defend. Although control group used more rejection comparing with treatment group, evidence based rejections' percentage over all rejections was higher in treatment group. In control group which were similar to traditional class, construction aspect of argumentation was emphasized more whereas evidence based critique aspect and movement in class were dominant in treatment group teachers' classes. In conclusion, findings suggested that if non-directionality combines with critique aspect, interaction and movement in class, it results with evidence based arguments. On the other hand, if directionality combines with construction aspect of argumentation, and non-movement with limited interaction, arguments without evidence are proposed.

In another study, Sampson and Clark (2011) proposed conditions in argumentation practice that increase quality of arguments. Five conditions were found that increase quality of arguments which are the number of ideas about content, how individuals respond ideas of each other, how often individuals challenge the ideas, the criteria individuals use to distinguish ideas, how group members used available data. Four groups of tyriads who are high school students participated in the study. This study identified the product as high quality written science argument and the task involved evaluating six alternatives for a discrepant event. This task provided students an opportunity to engage in scientific argumentation. Groups had homogeny in terms of gender (all girls or all boys). Regarding content knowledge, students completed the content knowledge test prior to the study and each group included students having different level of content knowledge. By this way, content knowledge effect on argumentation process was removed in this study. Prior to the study, scientific argument structure was taught to the students. Accordingly, scientific argument included three aspects which are explanation, evidence and reasoning. In this argument structure, explanation refers to claim that is the assertion; evidence is the data which is the fact that supports claim; and reasoning is the warrant that connects data and claim. After students engaged in argumentation in discrepant event, students were categorized based on their arguments (products of their argumentation process) as lower performers and higher performers. All the groups tried to solve task and all of them were active participants. Discussion for each group lasted 40 minutes to select best explanation.

Findings showed that the number of ideas related with content (i.e., number of discussed science topics in argumentation) was much more in higher performers' discussion than lower performers' discussion. Lower performers proposed fewer ideas about content to be discussed. Higher group mentioned 30 content related ideas; on the other hand, lower group mentioned 15 ideas. These ideas were related with both context, background knowledge, and other contexts. The more and diverse content related ideas result in the more discussion environment. Moreover, high performer group proposed more accurate ideas found in supplied materials, more accurate ideas not found in materials, inaccurate ideas found in materials, and inaccurate ideas not found in materials. When high performer group discussed deeper in topic, they proposed more ideas and when this group proposed more ideas, their argumentation included more inaccurate ideas including misconceptions comparing with low performer groups.

Second major finding was about how students respond to others. Higher performers tend to discuss all ideas before accepting or rejecting them. On the other hand, lower performers discussed less the ideas and willing to accept or reject ideas without further discussion. How group members responded to proposed ideas were coded based on four categories namely; accept (agreement no discussion), reject (disagreement no discussion), discuss (question, challenge, revise, further discussion), and ignore (no response). Coding of responses to proposed ideas showed that discuss rate in higher performers was 56% whereas discuss rate in lower performers was 43%. Moreover, accept and reject proportion for lower performers was more comparing with the higher performer group. Ignore proportion was similar for both groups. Higher group thought that each idea is product of cognition and need to be tested, evaluated and criticized. On the other hand, lower group did not discuss so much about ideas. As a result, lower performers' discussions were not in-depth. Ideas were end point in lower performers' discussion, but ideas were starting point for higher performers' discussion.

As third major finding, nature and function of the discussions were different. While higher performers challenged each other in discussion episodes, lower performers used challenge less and lower performers used seeking information, justifying own idea, and supporting others more in their discussions. Unit of analysis was conversational turns in discussion episodes. Four different categories emerged when comments about task in discussion episodes were analyzed which are information seeking (i.e. request more information and request clarification), exposition which is about people own ideas (i.e., proposing, clarifying, and justifying ideas), oppositional comments which is challenging the ideas of others, and co-construction of knowledge with others. Co-construction includes summarizing, revising, supporting, and adding ideas to ideas of others. This study showed that oppositional comment (i.e., criticizing each other) is beneficial for argumentation because students did not give up discussion when their ideas were criticized. Moreover, oppositional comment brings about in-depth discussion. While higher performer group used oppositional contribution more, lower performers used co-construction of knowledge and information seeking more. Expositional category was used similarly in both groups. Oppositional comment is characteristics of scientific argumentation. So, it is expected that higher performer groups used it more than lower performers did. Even though lower performers rejected the ideas more, they did not challenge these ideas. Lower performers focused on elaborating idea, asking question, agreeing and supporting. These students are not willing to discuss, critic, and challenge. Even though content is inaccurate, students did not change the nature of conversation and they did not challenge the ideas. Higher performers used more disagreement, challenge and critic. There was no polarization of viewpoints or students did not end discussion when they challenge each other. On the other hand, discussion continued further. Explanations, evidence and reasoning were criticized.

Forth finding of the study was about criteria that students used in argumentation. Accordingly, criteria that groups used differed when evaluating alternative viewpoints. Higher performers were careful in using criteria. They considered whether idea and evidence is consistent and whether idea and theory is consistent. On the other hand, lower performers used informal criteria mostly. They focused on whether explanation and background is consistent, or explanation and personal expectation is consistent. Higher performers used rigorous criteria (e.g., coherence of explanation) to justify and challenge ideas more than lower performers. On the other hand, lower performer group used informal criteria (e.g., appeal to authority) to justify and challenge ideas more than higher performers. The ones who used informal criteria frequently developed more misconceptions. High performers used rigorous criteria frequently, and so they identified and eliminated the plausible but wrong ideas.

Chen et al., (2016) claimed that it is possible to observe students' development regarding argumentation because classes had no argumentation experience in general. Researchers selected ecosystems and human body systems units to understand whether students transfer their argumentation skills to other contexts. The study lasted 16 weeks and 22 fifth grade students participated in the research. Class discussions were transcribed and researchers reached each participant's arguments revealed in argumentation process. Each individual's contribution was determined based on the number of their talking in discussions. Students' critics and contributions were coded based on social negotiation aspect of argumentation. Six different categories emerged as a result of social negotiation based analysis. These codes are information seeking (i.e., requesting more information about data source and procedure), elaboration (i.e., providing more information when question is asked to increase validity of claim, sufficiency of evidence and logic of evidence), challenging (i.e., testing procedure, accuracy of claim, quality of evidence, quality of reasoning, coherence between question and claim, coherence between claim and evidence), defending (i.e., trying to persuade speaker using simple defending and evidence based defending), supporting (i.e., thinking like speaker and using simple supporting and evidence based supporting) and rejecting (i.e., contradiction with speaker including ideas like claim does not fit with research question).

Findings of the study showed that students reached much more aspect of social negotiation over time. They participated in the discussions more as time progressed.

First rounds were based on simple questioning and answering. Information seeking and elaborating categories were used primarily. Debates and justification of claims were not used in first phases. In following phases, more critical elements of social negotiation which are challenging, rejecting, supporting and defending were put forward. In interviews, students reported that because of their lack of knowledge about argumentation, they did not participate in discussions. Statistically, both numbers of categories and argumentation increased as time went by. Dialogs between students evolved to whole class discussion over time. Students also reached scientifically acceptable knowledge through their explanations, their peers' critics, and students' answers to those critics. Also, students defending, supporting, rejecting ideas were based on evidence as time went on although these ideas were based on arguments without justifications at the beginning. This study showed that students can develop their oral arguments through practice. However, time and practice are needed for this development. They engage in argumentation process through practice and they construct and criticize the knowledge. At the beginning, students focused on construction of their arguments and ignore their peers' critics. Moreover, because of they did not understand their friends' arguments at the beginning; they focused on elaborating ideas and so challenging the ideas stayed in the background. While discussion process increased over time, students develop their challenges, rejections, defense, supports, critics and construction of knowledge. Students did not only criticize their peers, they also supported their peers' arguments to construct more complete arguments.

Kim and Song (2006) studied with 3 groups of middle school students. These 3 groups included 8 students in total. In this study, students first planned and conducted experiment, second wrote report and third defend their arguments to other groups. Students spent most of their time to decide their experiment topic and identify the problem. Then, they presented their written reports with other two groups to defend their own report and they criticized other two groups' reports when they were reported. Therefore, argumentation occurred between 2 groups as two rounds (i.e. as reporter and as opponent). As a result, three different discussions were conducted. Data was collected by discussion videos and student reports. Interviews followed this process. Via questionnaire, students explained their views about advantages and disadvantages of this inquiry activity. In data analysis, researchers found two different categories for students' argumentation strategies which were conflict and cooperative strategies. Regarding strategies students used in argumentation, researchers used cognitive and social aspects of argumentation. Argumentation is cognitive because it includes components of argument (e.g., warrant), but argumentation is social activity because it depends on communication and persuasion. Hence, cognitive strategies included questioning, elaboration, clarification, analogy, hypothesizing, and authorization. On the other hand, social strategies included conflict to control argumentation (repetition, cutting short, challenging, muttering) and cooperation to control argumentation (negotiation, explicit closing, suggestion).

The inquiry activity consisted of experiment and argumentation activity. After experiment, students wrote reports. Reports were written for peer review and discussion. These reports initiated reflective thinking and argumentation. After preparation of written arguments, stages of critical discussion start. Two activities which are argumentation and experiment provide feedback to each other. Argumentation activity gives feedback to the experiment. For example; argumentation result says that review the hypothesis. After this feedback, group reconsidered the hypothesis and conducted new experiment based on new hypothesis. Likewise, argumentation gives feedback about method and group changes the method of the experiment. On the other hand, experiment is prior to the discussion because claim and evidence of the argumentation are derived from experiments. This shows that argumentation and experiment are linked. Succession of experiment and argumentation revealed a circular structure of inquiry.

For whole argumentation process, researchers provided four stages and these stages formed different types of discussions. Before stages, students needed to decide on what to argue and they are supposed to explain the meaning of the content about what they discuss. Then, first stage "focusing" occurs. In focusing stage, uncertainties were eliminated. Students try to focus on their arguments in this stage. The aim is to form a common ground for discussion. At the end of focusing stage, students resolved ambiguity. Second phase is exchanging in which reporter provides information about the detail of their research. The aim is not to discuss in this stage. After, topic of presenter group is understood, debating stage starts. Opponent’s answer to reporter determines debating which are confrontations (i.e., rejecting reporter), coexistence (i.e., supporting reporter but also supporting other arguments) and consensus. After opponents’ reaction, reporter answers and these result in three different conditions which are ignoring argument, accepting argument, and reaching an agreement. Three different closings can follow debating which are explicit closing (i.e., argumentation ends with agreement explicitly), implicit closing (i.e., Argumentation ends with agreement implicitly. For example, summary or repetition is used at the end), and circumstantial closing (i.e., argumentation ends with conflict and there is no resolution. Closing can include silence and muttering.). Although researchers found different types of discussions (e.g., focus-exchange-confrontation-reaching agreement-explicit closing), researchers did not report that one type of discussion is better than another in their explanation. Researchers claimed that focusing is the most important stage of discussion. Teachers should inform students about the process, concepts and other details which provide ground for communication during argumentation.

At this point it is important to note that, argumentation studies focusing on argumentation process is not unique to argument-based inquiry. Researchers also focused on argument driven inquiry to reveal students’ engagement in argumentation process. For example, in an argument-driven inquiry (ADI) study which is another immersion approach for argumentation, Walker and Sampson (2013) examined how students’ argument performance, oral argumentation and written argumentation changed over time. A total of 46 students enrolled in chemistry lab course participated in the study. Study included two sections and each section included six groups comprised of four individuals. At the beginning of the course; ADI handout was provided to students. Handout included some information about topic, statement of the problem, research question, available material, precautions, starting suggestions. Students were supposed to report what they did in the class. Experiments were ranked from the simplest to the most complex one. Use of lab and steps of the ADI were introduced to the students. A total of seven chemistry investigations were conducted, but five of them were reported. TAP was not suitable for analysis of classroom reports or oral argumentation because it ignores the interaction between individuals. Moreover, its components were not separated easily in analysis. Hence, video analysis was analyzed based on Assessment of Scientific Argumentation in the Classroom (ASAC). ASAC includes argumentation observation protocol consisting of cognitive, epistemic and social structure. Cognitive structure focused on the ideas about different claims, their reasons and rationalization of ideas. Epistemic structure was based on supporting and challenging claims. On the other hand, social structure was based on the interaction between participants. All groups were observed separately. Findings of the study suggested that students' responses to the task performance developed over time. Their rationale and claim developed over time whereas evidence did not improve. On the other hand, previous studies suggested that students have difficulty in constructing rationale because rationale develops through practice. Therefore, researchers suggested that practicing argumentation is better than learning structures of argument for students when students construct rationale. Moreover, increasing content knowledge and ADI experience assisted to produce better rationale. So, performance was highest at the end of the program. Second finding was related with oral argumentation. Quality of oral arguments varied from one topic to another. In early investigations, students' argument generations were better than argument evaluation. However, later investigation results showed that students' argument evaluation was better than argument generation. Students shared the responsibility by establishing pre-set roles, so they did not make equal contribution on argument generation process which did not result in further improvement of argument generation. On the other hand, argument evaluation was new for students, and they equally evaluated the arguments showing increase in argument evaluation quality. In first investigations, evaluations of ideas were held at the end of the investigation and students left the lab in these weeks early. Thus, argument evaluation was lower than argument generation in these weeks. In conclusion, this study showed that use of ADI improved both quality of arguments and argumentation process. While students improved their claim and rationale in their arguments, they could not develop evidence in their arguments. Regarding argumentation process, argument evaluation aspect that is criticizing arguments developed in last weeks because it was new at the beginning of the study and students could not adapt it. On the other hand, students could not develop argument generation that is constructing arguments so much because pre-set rules giving responsibility to students inhibited them to engage in argument generation process equally.

### 2.6.2. Summary of Studies focusing on Argumentation Process.

In conclusion, all research on students' engagement in argumentation process was qualitative and each study informs about nature of students' argumentation. Cavagnetto et al. (2010) study showed that middle school students mainly made effort to engage in argumentation, but these students could not provide rebuttals that decrease the quality of argument. Likewise, while students constructed arguments, they could not criticize each other. However, students were able to transfer their engagement in argumentation from one topic to another supporting the view that topic is not so much important in argumentation. In second study, Choi et al. (2014) study also showed that students actively engaged in argumentation. As in the Cavagnetto et al. (2010) study, students' participation was similar in different contexts. In this study, students focused on evidence rather than claims. The study also suggested that type of inquiry is important for argumentation process. Accordingly, this study showed that students focus on experimental design if they prepare their research question. On the other hand, if research question is directly presented to students, students focus on resources of evidence during argumentation. In third study, Kim and Hand (2015) focused on the conditions that result in better arguments. Researchers claimed that the ones who criticize more can produce better arguments. Moreover, if teachers do not give direction to students, if teachers move in class and interact with students more, students produce high quality arguments. Likewise, when these conditions are provided, Kim and Hand (2015) claimed students used more evidence based reasoning. In forth study, Sampson and Clark (2011) compared argumentation process of students having high quality arguments and students having low quality arguments at the end of argumentation process.

Researchers claimed that students who had better arguments used more number of content knowledge, and discussed on the ideas more than students having low quality argument. Likewise, students having better arguments used oppositional comments more than exposition, information seeking and co-construction of knowledge components of discussion. Moreover, students producing high quality arguments used rigorous criteria to evaluate claims like scientists, but students producing low quality arguments used informal criteria. In another study, consistent with other research (Cavagnetto et al., 2010; Choi et al., 2014), Chen et al. (2016) claimed that middle school students transfer their argumentation skills to different contexts. Chen et al. (2016) study showed that students' engagement in argumentation process improved over time. Likewise, students start to use more evidence in their arguments over time. Students used information seeking and elaboration components more in first weeks, but they used challenge, reject, support and defend aspects of oral argumentation in following weeks. Similarly, students focused on constructing arguments in early weeks, but they focused on both constructing and criticizing arguments in following weeks. Kim and Song (2006), on the other hand, identified stages that middle school students followed in argumentation which are focus, exchanging ideas, opponent reaction, reporter answer and closing. Researchers claimed that use of these stages form discussion types happening in class. According to researchers, these discussion types are exchanging information, consensus, coexistence, confrontation, and extension. Finally, in their ADI research, Walker and Sampson (2013) reported that students develop their rationale and claim over time, but the use of evidence did not change. Students' familiarity to ADI approach and increasing content knowledge also facilitated students' use of rationale that is one component showing argument quality. As observed in Chen et al.'s (2016) study, students focused on generating arguments in first weeks, but evaluation of arguments got more attention than generating arguments in following weeks. In the next part, argumentation studies focusing on argumentation schemes are presented.

#### 2.6.3. Reflections from Literature to Current Study regarding Argumentation Process.

Next, how previous literature contributed on current research about argumentation process is explained. First of all, vision-1 approach of scientific literacy was adopted in this study. According to this idea, variables which are directly related with scientific practice such as content knowledge, epistemological beliefs and science process skills are considered. When literature was reviewed regarding argumentation studies, review results showed that only immersion argumentation orientation (see Cavagnetto, 2010) fits with vision 1 approach of scientific literacy. Because of this reason, immersion was adopted as argumentation orientation in this study before this treatment was prepared. Literature review about immersion argumentation orientation also showed that researchers mainly used two different immersion approaches in their treatments which are argument-driven inquiry (ADI) (e.g., Walker & Sampson; 2013) and argument-based inquiry (ABI) (e.g. Hand et al., 2004). Due to fact that ADI was designed to reveal undergraduate students' argumentation in laboratories, ADI was not used as treatment in current study. On the other hand, ABI has been used in different grade levels like elementary (e.g. Taylor et al., 2018), middle school (e.g. Günel et al., 2012), high school (Kingir et al., 2012) and undergraduate (Cronje et al., 2013), current study used ABI as argumentation treatment. While lesson plans, teacher training, student training, and pilot study were prepared, ABI content was considered. For example; asking questions, doing investigations, obtaining data, using evidence, proposing claims, and making discussions were focus points in all activities. All plans were constructed on this approach and students' engagement in argumentation process was examined based on the content of the ABI approach. If another approach had been adopted in this study, most probably students' engagement in argumentation process would have been different. Although ABI approach or immersion do not include explicit teaching of components of argument (e.g., justification), components of argument and their examples were told to the students when students were trained prior to the study. By this way, students got familiarity to the study. Review of epistemology and argumentation studies also provided information about students' engagement in argumentation process. These studies frequently reported that there are two types of data that participants used when they engaged in argumentation which are first hand and second hand data. In line with this, these two different data types were used in current study to reveal how different types of data affect students' engagement in argumentation process. Furthermore, previous research also informed about data collection process of current study. Previous research (e.g., Kim & Song, 2006) showed that researchers used video records to examine students' engagement in argumentation process. In line with this, class discussions were recorded by video cameras throughout the study. Next, data analysis for students' engagement in argumentation process was fed by existent literature. Accordingly, Sampson and Clark's (2011) analysis of nature and functions of discussions fits with the data obtained from this study. Therefore, Sampson and Clark's (2011) analysis was used when students' engagement in argumentation was analyzed. Interestingly, Chen et al. (2016) proposed data analysis for students' engagement in argumentation in ABI treatment, but this analysis did not fit with the data obtained in this study because this analysis did not include expositional comments (i.e., the comments that students propose about their ideas); however, expositional comments were common in current study and therefore, Chen et al.'s (2016) analysis were not used in current study. Likewise, unit of analysis to analyze students' engagement in argumentation was borrowed from Sampson and Clark's (2011) study. According to this, unit of analysis was determined as conversational turns in discussion episodes. Finally, previous studies (e.g., Chen et al., 2016) showed that when researchers analyzed participants' engagement in argumentation process, they first used content analysis and after that they used descriptive statistics such as frequency analysis to see changes of participants' performance throughout the process. In line with this, in this study, first Sampson and Clark's (2011) coding was used in data analysis process as content analysis and after that descriptive statistics were used to compare participants' engagement in argumentation process throughout the study.

All in all, this literature review included five different main themes regarding argumentation studies. First part included theoretical consistency between argumentation and content knowledge, and practical research on argumentation impact on content knowledge. Likewise, second part included the theoretical connections between argumentation and epistemological beliefs, and practical work about argumentation studies dealing with epistemological beliefs. Third part was about contribution of argumentation on science process skills. Therefore, theoretical match of argumentation and science process skills and practical work regarding argumentation impact on science process skills were reviewed in third part. After that, argumentation studies examining argumentation schemes were presented and finally, argumentation studies dealing with students’ engagement in argumentation process were reviewed. Methodology of the research is given in next chapter.

# CHAPTER 3

# METHODOLOGY

Current research examined the effect of argumentation based inquiry approach (ABI) on $6^{th}$ grade level students' content knowledge (e.g., matter and heat), epistemological beliefs, and science process skills. Moreover, this study aims to reveal $6^{th}$ grade students' engagement in argumentation process and their argumentation schemes used in whole class discussions. In this chapter; research design, sample and participants, variables, process prior to the data collection, data collection, verification of independent variable, data analysis, validity, ethics, limitations and assumptions are presented.

## 3.1. Research Design

Research questions determine the type of research design. While quantitative research design is used to answer research questions that explore relationship between variables and look for causes of relationships, qualitative research design is used to answer research questions that explain the natural phenomena (Fraenkel, Wallen, & Hyun, 2012). In this study, the impact of argumentation based inquiry approach on selected variables is explored in first three research questions which are "Is there a change in $6^{th}$ grade students' content knowledge from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?", "Is there a change in $6^{th}$ grade students' epistemological beliefs from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?and "Is there a change in $6^{th}$ grade students' science process skills from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?. Therefore, this study has quantitative research design on the one side. However, last two research questions of the study are based on explaining students' engagement in argumentation process and their argumentation schemes. Hence, this study has qualitative aspects too.

Although this study has quantitative methods and qualitative methods, data obtained from these two methods were not mixed explicitly. Different data obtained from different methods were used to answer their corresponding research questions. In this point, Creswell (2006) claimed that “Mixing is the explicit relating of the two data sets. A study that includes both quantitative and qualitative methods without explicitly mixing the data derived from each is simply a collection of multiple methods.”(p.83). Due to fact that current study did not mix data, this study can be called as a collection of multiple methods.

In the quantitative part of this research, one group pre-test post-test design was employed. More specifically, there were four experimental groups (classes) taught by argument-based inquiry approach, but there was no control group in this study. Prior to the treatment, all these four groups were administered content knowledge tests for matter and heat unit and electricity unit, epistemological belief questionnaire and science process skills test as pre-tests (Time 1). Then, the ABI was implemented in the unit of matter and heat as treatment. Following this unit, content knowledge test for matter and heat unit was again administered as post-test (Time 2). Next, students received the ABI instruction in the unit of electricity.  After the treatment is completed, groups were administered the content knowledge tests for electricity unit, epistemological belief questionnaire and science process skills test as post-tests (Time 3). Research design of the quantitative part is given in Table 3.1.

Table 3. 1

*Quantitative Part Research Design*

| **Groups** | **Time 1** | **Treatment** | **Time 2** | **Treatment** | **Time 3** |
|---|---|---|---|---|---|
| Experiment Groups | Content Knowledge test for Matter and Heat Unit (pre-test)<br><br>Content Knowledge test for Electricity Unit (pre-test)<br><br>Epistemological Belief Questionnaire (pre-test)<br><br>Science Process Skills Test (pre-test) | Argument based inquiry (ABI) treatment | Content Knowledge test for Matter and Heat Unit (post-test) | Argument based inquiry (ABI) treatment | Content Knowledge test for Electricity Unit (post-test)<br><br>Epistemological Belief Questionnaire (post-test)<br><br>Science Process Skills Test (post-test) |

On the other hand, in qualitative part of this research, case study was adopted. A case is bounded by time and activity (Creswell, 2009). In this study, ABI was used as treatment and how students experience this treatment was observed in qualitative part. Due to fact that this study explores students’ experience to ABI treatment, ABI treatment was selected as specific case for the qualitative part of the study. ABI treatment as case also determined its boundaries. For example; activities, investigations and discussions that students engaged are results of ABI treatment. Two classes of students participated in qualitative part of this study and these two classes are defined as sub-units of this case study. Figure 3.1 summarizes research questions, corresponding research designs, and type of research.

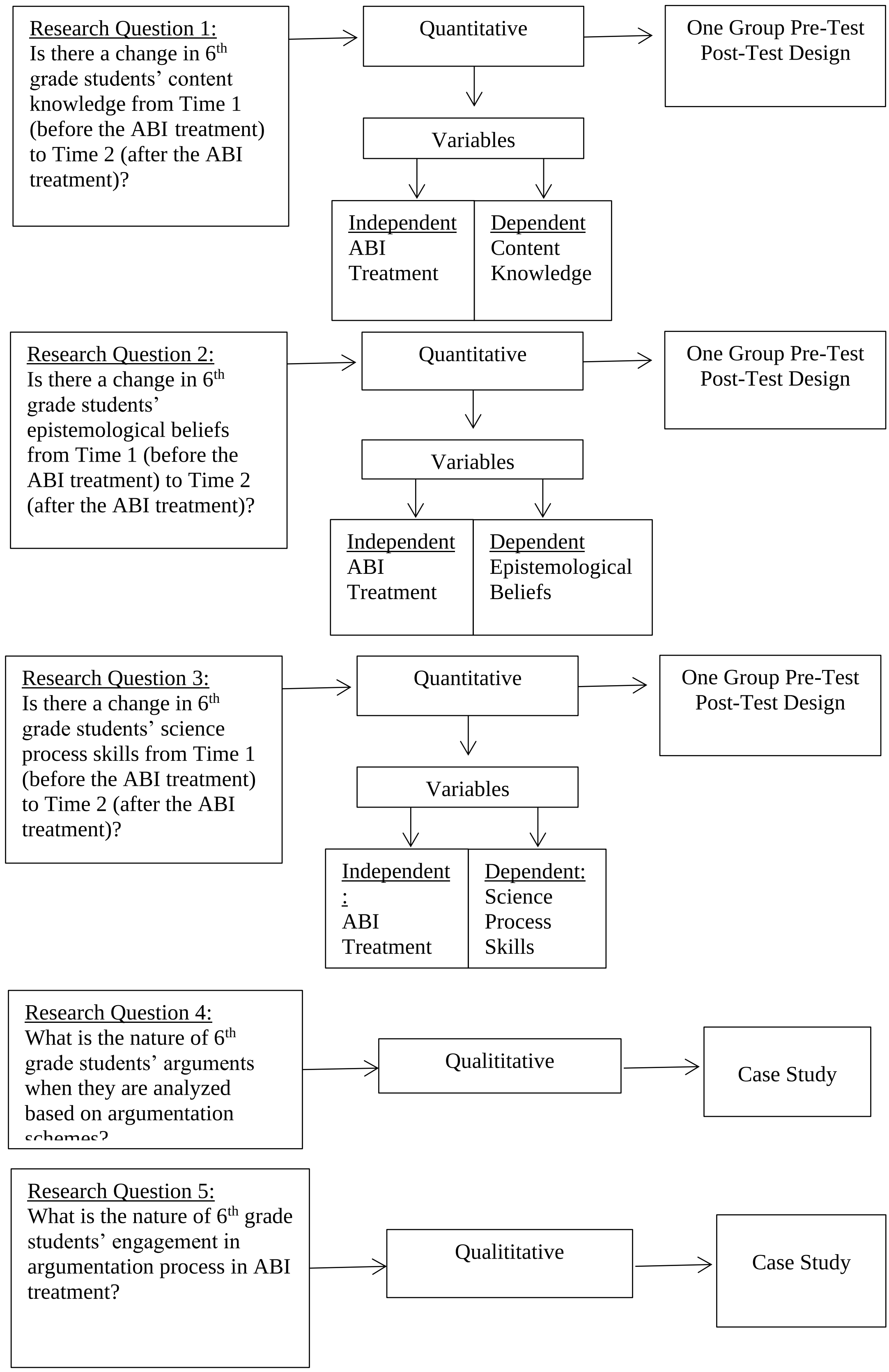



*Figure 3. 1* Summary of research questions, research designs and type of research

### 3.2. Participants

Regarding quantitative aspect, all 6$^{th}$ grade students enrolled in public schools in Ankara form the target population of the study. On the other hand, accessible population is all 6$^{th}$ grade students enrolled in public schools in Çankaya district. Convenient sampling strategy was used to select the sample from the accessible population. More specifically, one middle school located in centre of Çankaya district which was easily accessible to the researcher was selected due to time, energy, and cost considerations. A total of four classes taught by the same teacher involved in the study. All these classes received ABI treatment. Number of students in each class was almost the same ranging from 17 to 18. Therefore, a total of 71 (36 males and 35 females) Grade 6 middle school students constituted the sample of the quantitative part.  Students' ranged in age from 11 to 12. Students' previous term science grades were between 40.33 (minimum score) and 98.22 (maximum score) over 100. The mean scores of participants' previous term science grades were 78.61 with a standard deviation of 13.48. Based on their previous term science scores, it can be claimed that students are not bad in science lessons. All classes were comparable in terms of age, previous science grade, and gender distribution. Prior to the study, each class followed the same curriculum suggested by MONE (2013). Although the school had science laboratory, the teachers mainly did not use this laboratory. Both pilot and main studies were conducted in this laboratory.

In qualitative part, two classes were randomly selected. Each of the class was sub-units of case in this study. Data were collected and analyzed separately for each sub-unit. After that findings of two sub-units (i.e., class) were compared and contrasted. Accordingly, a total of 35 students (17 students from class 1, 18 students from class 2) participated in this part. In class 1, there were 9 girls and 8 boys. The mean of the previous term science grades of participants was 77.33 for class 1. On the other hand, class 2 included 8 girls and 10 boys. Class 2's science score mean was 73.76 in the previous term.

In laboratory, there were four desks prepared for group work. Students were separated in four groups for each class. Each group consisted of 4-5 students. In pilot study conducted in spring semester of 2018, groups were heterogeneous and teacher formed the groups. However, some groups did not work well because some students did not want to work with others and there were some management problems which can cause threats to implementation of the ABI. Therefore, students were allowed to select their group members in main study. This strategy was useful because students could better focus on their experiments and discussions in main study. Likewise, there was less management problems throughout the study comparing with pilot study. Groups were still heterogeneous in terms of science achievement, personality, gender although students formed the groups. Main study was conducted in 2018-2019 spring semester.

Each of the participants was labelled based on their class, desk number, and position because of ethical concerns. Hence, each participant's label included three digits. For example students in first class, first desk and first position was labelled as student 111; on the other hand, student in second class, third desk and fifth position was labelled as student 235. Laboratory seating arrangement of class 1 and class 2 are presented in figure 3.2 and figure 3.3.

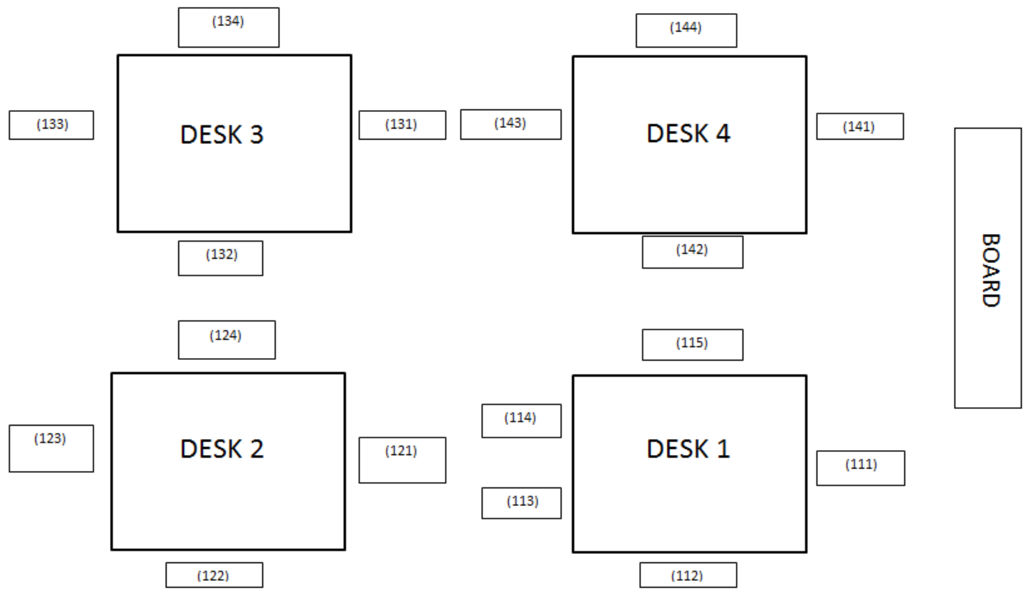



*Figure 3. 2* Laboratory seating arrangement of class 1

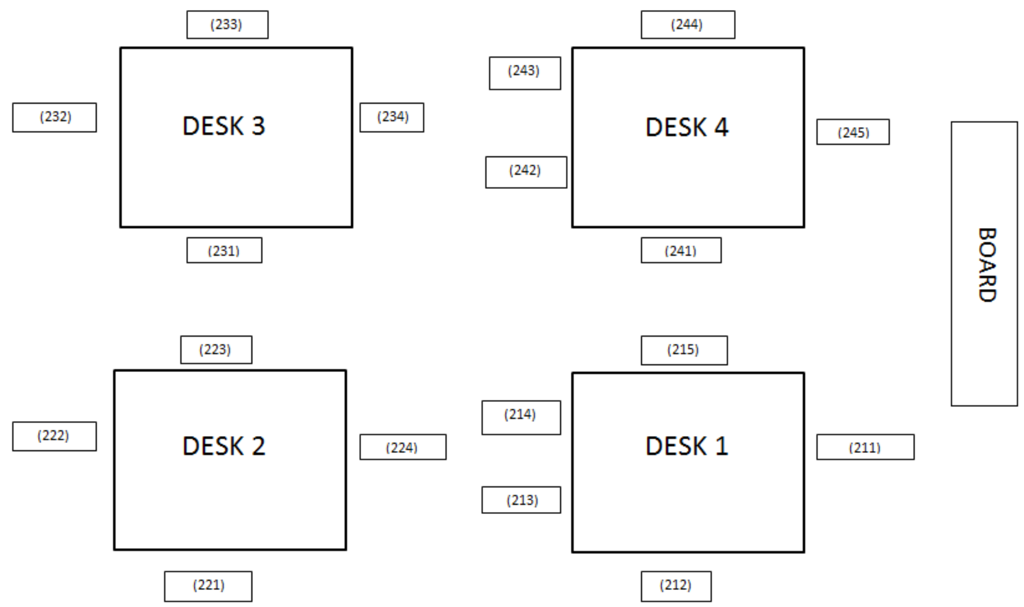



*Figure 3. 3* Laboratory setting arrangement of class 2

## 3.3. Variables

In quantitative part, the variables can be categorized as independent variable and dependent variables. Independent variable that researcher manipulated in this study is the use of argument-based inquiry treatment. All participating classes took this treatment, so all classes were experiment group. There was no control group in this study. There were three dependent variables which are participants' content knowledge, epistemological beliefs, and science process skills.

In qualitative part on the other hand, two variables were examined throughout the study including students' engagement in argumentation process and their argumentation schemes.

## 3.4. Process Prior to the Data Collection

Preparation process was done prior to collecting main data. Preparation process included topic selection, lesson plan preparation process, school selection, teacher training, and pilot study. This preparation process started in May 2016 and finished in May 2018 when pilot study completed.

### 3.4.1. Topic Selection.

In Turkey, it is mandatory to follow curriculum suggested by MONE (2013). Therefore, current study had to consider curricular objectives. Although curriculum suggested teachers to use student centered approaches like argumentation, curricular objectives hardly mention the components of arguments (Cetin et al., 2016). Therefore, middle school science curriculum objectives were analyzed from grade 5$^{th}$ to 8$^{th}$ at first. Science units which have objectives including experiments and discussions were at the centre of topic selection because it was thought that objectives including experiments and discussions can better align with argument based-inquiry approach.

Previous research reported that duration of argumentation treatments is important for efficiency of treatment. These studies suggested that if duration of treatment increases, students benefit more (Hong, Lin, Wang, Chen, & Yang, 2013). Therefore, two different science units were selected in order to increase the duration of the treatment. The selected units were Matter and Heat unit and Electricity in $6^{th}$ grade curriculum. Matter and Heat unit included more objectives than Electricity unit. Therefore, teaching of matter and heat unit takes more time than teaching electricity. Core ideas of Matter and Heat unit are heat conductors, heat insulators, thermal insulating products, and types of fuels. On the other hand, in electricity unit, core ideas are electrical conductivity, and factors affecting bulb brightness. To sum up, this study had six core ideas. Six lesson plans were prepared based on core ideas. Teaching of each core idea using ABI treatment lasted one week; hence, treatment lasted six weeks both in pilot and main study. Next part informs about lesson plan preparation process.

### 3.4.2. Lesson Plan Preparation Process.

Argument-based inquiry implementation requires the use of two templates which are student template (see Appendix A) and teacher template (see Appendix B). While student template is a scaffolding used by students to construct their arguments, teacher template is used as scaffold by teacher to conduct phases of argument-based inquiry (Hand & Keys, 1999). Therefore, teacher template was considered while preparing lesson plans. Each lesson plan format was consistent with the structure of the teacher template. Accordingly, lesson plans started with curricular objectives that are covered in that week. Then, the lesson plans continued with the following phases: eliciting students' prior knowledge, activities prior to experiments, engaging in laboratory activities, negotiation phase-1 (individual negotiation), negotiation phase-2 (small group discussion), negotiation phase-3 (whole class discussion), negotiation phase-4 (reflection), and evaluation. In other word, each lesson plan was a teacher template that was prepared to assist teacher's effective implementation of ABI.

The lesson plans prepared for core ideas of heat conductor, heat insulator, electrical conductivity and the factors affecting bulb brightness included all of the above mentioned steps. All these core ideas included experiments and this situation fits with the ABI implementation. One of the lesson plans prepared considering these core ideas are given in Appendix C. On the other hand, core ideas which are thermal insulating products and types of fuels did not include any experiments. Therefore, some adjustments on lesson plans (teacher templates) about these core ideas were done. Accordingly, lesson plans started with objectives as usual, and eliciting students' prior knowledge, negotiation phase-1 (individual negotiation), negotiation phase-2 (small group discussion), negotiation phase-3 (whole class discussion), negotiation phase-4 (reflection), and evaluation were same. However, phase-2 and phase-3 which are activities prior to the experiment and engaging in laboratory activities were replaced with prior activities and engaging in activities in these two core ideas' lesson plans (i.e., thermal insulating products and types of fuels) where making experiment was not possible. One of these two lesson plans which were prepared to teach core ideas that do not include science experiments are given in Appendix D. While students collected their own data in other four core ideas' activities where they conducted experiments, they did not collect data in thermal insulating products and types of fuels topics. Students constructed and evaluated their arguments based on the evidence cards prepared and given by researcher in these core ideas. Therefore, it can be said that students used first-hand data in heat conductors, heat insulators, electrical conductivity and factors affecting bulb brightness topics where they conducted experiments, and they used second-hand data in thermal insulating products and types of fuels topics where evidence cards were provided to them. Table 3.2 summarizes the lesson plans' core ideas, their application weeks, sequence of steps followed in ABI instruction and type of data planned to be used by students in given lesson plan.

Table 3. 2

*Summary of Lesson Plans*

| Weeks | Lesson Plans | Core Ideas | Steps of Lesson Plans | Type of Data |
|---|---|---|---|---|
| Week 1 | Lesson Plan 1 | Heat Conductors | Eliciting students' prior knowledge, activities prior to the experiment, engaging in laboratory activities, negotiation phases 1-2-3-4, evaluation | First-hand |
| Week 2 | Lesson Plan 2 | Heat Insulators | Eliciting students' prior knowledge, activities prior to the experiment, engaging in laboratory activities, negotiation phases 1-2-3-4, evaluation | First-hand |
| Week 3 | Lesson Plan 3 | Thermal Insulating Products | Eliciting students' prior knowledge, prior activities, engaging in activities negotiation phases 1-2-3-4, evaluation | Second-hand |
| Week 4 | Lesson Plan 4 | Types Of Fuels | Eliciting students' prior knowledge, prior activities, engaging in activities negotiation phases 1-2-3-4, evaluation | Second-hand |
| Week 5 | Lesson Plan 5 | Electrical Conductivity | Eliciting students' prior knowledge, activities prior to the experiment, engaging in laboratory activities, negotiation phases 1-2-3-4, evaluation | First-hand |
| Week 6 | Lesson Plan 6 | Factors Affecting Bulb Brightness | Eliciting students' prior knowledge, activities prior to the experiment, engaging in laboratory activities, negotiation phases 1-2-3-4, evaluation | First-hand |

After lesson plans were prepared, they were sent to two science education researchers having expertise in ABI research. Researchers' feedbacks on lesson plans were mainly related with type of inquiry. Experts thought that there were so much direction in lesson plans, and students should get more independence in ABI research. For example, students should decide the materials they would use. Lesson plans were adjusted based on experts' recommendations and lesson plans included less direction and provided more independence to students in latest form. By this way, type of inquiry changed in lesson plans from closed inquiry where students have less independence to guided inquiry where students have more independence. This was also suitable for requirements of MONE (2013) because MONE (2013) asks teachers to use guided inquiry in 6th grade level.

Preparing lesson plans had three main benefits to the current study. First, it prepared researcher to the study. Researcher got familiarity to the phases of teacher template because lesson plans addressed phases of teacher templates, and was aware of the important and difficult parts of ABI instruction. Secondly, lesson plans were used as part of teacher training. The teacher was no knowledgeable about ABI instruction prior to the study. Therefore, lesson plans were shared with teacher and teacher implemented these lesson plans in pilot and main studies. By this way, teacher got familiarity to the ABI approach. The last benefit of lesson plans might be about validity of current study. The lesson plans were prepared considering the phases of teacher template. Therefore, it is possible that lesson plans assisted teacher to implement ABI treatment effectively. By this way, implementation of lesson plans might generally contribute to the validity of study.

### 3.4.3. Teacher Training.

After lesson plans were prepared, researcher sought for a volunteer teacher in the selected school. After obtaining ethical committee approval, researcher went to school that study was held and had a conversation with school principals. School principals supported the study and directed researcher to the science teachers. There were two science teachers who were teaching in 6th grade levels. In this meeting with teachers, teachers reported their teaching orientations. While one teacher reported he mostly used direct instruction and questioning, other teacher reported she used questioning and discussion in science teaching. The teacher using discussion was considered to be more suitable for this study's context (i.e., ABI instruction) and it would take less time to train this teacher. Therefore, she was selected as the teacher to conduct ABI in her four $6^{th}$ grade level students.

The teacher was female and thirty five years old having ten years of teaching experience. She was graduated from elementary science education department of a public university. After working in three years in rural regions of Turkey, teacher started to work in this school. She had seven years of teaching experience in this school, so she had enough knowledge about school contexts. Before pilot study, researcher spent time with teacher in order to understand her orientation towards teaching and some of her real teaching was observed by researcher. These observations revealed that although the teacher asked questions and lead discussions in her classes, she was the authority and she mainly held teacher centered ideas in teaching. Therefore, it can be claimed that implementing teacher did not have student centered or constructivist ideas for teaching and learning. This situation could be problem for research. Moreover, the teacher did not know the ABI approach. Therefore, she needed training for effective implementation of the ABI.

Teacher training for ABI implementation included six meetings. Three graduate students who graduated from elementary science education department also participated in meetings. These students also assisted the implementation of ABI instruction both in pilot and main studies. The meetings were held prior to the pilot study on January 2018. The content of the meetings was science, science process skills, nature of science, argumentation, argumentation based inquiry and assessment of scientific argumentation in class.

At the beginning of our conversations, teacher focused on the product of science (content knowledge) and her main aim was to teach content knowledge emphasized in the curriculum. She had some naïve ideas about science. For example, she thought that science is certain; scientists follow one scientific method in their research. She also added that she usually did not use laboratory. It was inferred that the teacher might not know the construction of scientific knowledge because she did not use laboratory in science lessons. In line with this, the term "science" was discussed with teacher in the first meeting. In this discussion period, ways to obtain knowledge such as sense organs, logic and science are compared. Why science is better than other ways of obtaining knowledge was explained. In a similar vein, how scientists construct the knowledge was told with teacher. Therefore, common points of scientific research which are proposing problem, planning or designing a solution, and interpreting data were discussed with teacher. Then, there was a debate regarding how people see science. Therefore, positivist and post-positivist ideas were discussed.

Second meeting's topic was science process skills. Actually, the teacher was knowledgeable about science process skills, but these skills were reviewed because students were expected to use these skills when they conduct experiment and use laboratory materials. This review might assist teacher to focus on process skills when students make their experiments. Accordingly, basic process skills which are observing, inferring, measuring, communicating, classifying, and predicting and relationships between these skills were discussed at first. Then, integrated process skills which are defining operationally, formulating hypothesis, interpreting data, designing investigation, modelling, and presenting information were discussed. Teacher reported that she did not know operational definition; therefore, she found this review informative for her process skills.

In third meeting, conversation topic was epistemological beliefs because sophisticated epistemological beliefs support argumentation (Liu & Roehrig, 2019). In line with this, sophisticated epistemological beliefs and naïve epistemological beliefs' examples were shared with teacher. Accordingly, absolutist, multiplist, and evaluativist beliefs were shared with teacher and why evaluativist beliefs are hierarchically better than other two views to engage in argumentation was told in this meeting. Moreover, role of evidence in science, development, and certainty aspects of science was discussed with teacher. After these discussions, teacher claimed that she was surprised when she learnt that scientific theories do not transform to scientific laws, scientific laws are not proven, laws have limitations, and theories and laws are different kind of entities. Therefore, it is thought that this discussion was helpful for the teacher to improve her epistemological beliefs.

After talking about science, science process skills and nature of science, forth discussion continued with argumentation. In this conversation, components of argument which are evidence, warrants, rebuttals, and claim were explained initially. Next, two texts were provided to teacher about hunting and energy sources and teacher was asked to find out argument components in these texts. After that advantages of engaging in argumentation process for students were told to teacher.

After teacher became knowledgeable about argument and argumentation, argument-based inquiry treatment was presented to her in fifth conversation. Both teacher template and student template were explained to the teacher. While these templates were being told to teacher, a lesson plan about nervous system including teacher and student template was shown to teacher. It is expected that this lesson plan was useful for her understandings about argument-based inquiry and its implementation.

In last conversation, there was a conversation about to what extent teachers can be sure that they do argumentation in class. Therefore, characteristics of a class where ABI is used were explained to the teacher. In line with this, Assessment of Scientific Argumentation in the Classroom (ASAC) prepared by Sampson et al., (2012) was presented to teacher. Nineteen items including conceptual, epistemic and social aspects of argumentation were shown to teacher and teacher was asked to consider these items in her implementation of ABI research. By this way, quality of implementation was aimed to increase.

Teacher training was not limited with face to face conversation. Lesson plans were shared with teacher and teacher studied lesson plans before she implemented them each week. By this way, she became more familiar to the ABI instruction. Moreover, teacher implemented the ABI instruction in pilot study which lasted 6 weeks. It is expected that this practice also assisted her improvement. The details of pilot study are presented in next part.

### 3.4.4. Pilot Study.

Pilot study was conducted in 2017-2018 spring semester and it lasted 6 weeks. All the procedure was same with main study. Accordingly, four different classes participated in the study. Each class included four groups having 4-5 students. A total of 79 students participated in the pilot study. Three of the groups were selected randomly, and these groups' small group discussions were recorded. Moreover, two of the classes were video-recorded for whole class discussions. Prior to the pilot study, researcher met with each class and students were trained as their teacher. This training lasted four hours. In these training sessions, students learnt about inquiry, scientific debate, science, scientists, learning through inquiry, science process skills, and prompts of the ABI (i.e., student template). Training continued with demonstration about density. In this activity, students predicted which objects sinks and which objects float in water. After that, they observe the sinking and floating objects. Then, they explained their ideas regarding observations and they discussed with their peers. After this activity, researcher explained students that they made inquiry and engaged in argumentation process because students asked questions, proposed claims, used evidence, supported and defended their arguments. This activity also encouraged students to think that they are capable of making inquiry and engaging argumentation process. Application of the pilot study provided feedbacks for the main study. Some notes were obtained in each week and these notes were used in main study to improve quality of the treatment. These notes obtained in pilot study are as follows:

First week Notes (Heat Conductors Topic); 04.04.18

- Students' initial ideas show that students had prior knowledge about heat conduction because their predictions were correct. However, they could not make theoretical explanations for their predictions.

- While some classes had difficulty in heat conduction experiment, some classes performed well in their experiments.

- Solid oil, chick pea, hot water, and different kinds of spoons were provided to groups regarding heat conduction experiment. However, most groups did not prefer to use chickpeas in their experiments. Therefore, this material may not be given them in main study.

- Eliciting prior knowledge and designing of investigation lasted one hour.

- Conducting experiments, recording observations and constructing group claims lasted one hour.

- Whole class discussion lasted one hour.

- After whole class discussion, theoretical part was read from textbook and teacher informed students further about theoretical explanations regarding why some materials are better conductors than others. Teacher told that some materials' particles are in order and these particles are closer to each other comparing with other materials' particles in micro level. These order and close distance between particles make them easily conduct heat. This process lasted one hour too.

- Four hours were spent in total for heat conduction.

- Due to fact that time is limited, planned reading texts and videos were not used.

- Students were voluntarily to discuss, criticize, suggest and ask questions.

- Students have difficulty in writing their reports.

- Students need more directions because they could not connect their experiments and curricular objectives.

- Some students do not discuss and conduct experiments. Teacher tries to engage these students in process by asking questions and inviting them to ask questions.

- No technology related problem was encountered.

- Existence of video camera did not make students uncomfortable.

Based on the aforementioned first weeks' notes of pilot study, first week of the main study was designed. Accordingly, chickpeas were removed from the experiment for main study because students generally did not use chickpeas in their heat conduction experiment. Duration of the activities for the main study was also determined considering the findings of pilot study. In line with pilot study, one hour was given for eliciting prior knowledge and designing experiment, one hour was given for conduction of experiments and preparing group arguments, one hour was given for whole class discussion, and one hour was given for review of the topic and completion of missing points for the main study. Planned readings and videos were not used in main study because time was not sufficient to use them in pilot study. Other parts of the lesson that worked in pilot study remained the same and they were used in the main study as it was planned.

Second week Notes (Heat Insulators Topic); 11.04.18

- Students' prior knowledge about heat insulation was examined through the activity about a snowman wearing coat. Students were expected to explain in which condition snowman melts earlier. In first condition, snowman did not wear a coat, and snowman wore a coat in second condition. Most of the students thought that snowman that did not wear coat melt earlier and therefore, their content knowledge about heat insulation was not proficient in this prior activity. This discussion and designing experiment lasted one hour.

- In their experiments, students mainly compared the heat losses of water in different cups. In these experiments, cups were used as insulator materials. If the materials are better insulators, the heat loss of the water will be less. Conducting experiments and preparing group claims lasted one hour.

- Whole class discussion lasted one hour.

- Teacher informed students about theoretical explanations of heat insulations. Students read textbooks at the end. This process lasted one hour too. In conclusion, heat insulation activity lasted four hours in total.

- Students brake thermometers in their experiments in one of the classes. I should bring more thermometers in main study.

- Students did not use beaker when they collect data. On the other hand, they used half glass of water. In other word, they did not use ml in their measurement. Therefore, their measurement may not be accurate.

- I need to answer the following question: Is there any effect of initial temperature of water on heat loss? (After I asked this question to myself, I learnt that initial temperature was important in heat loss. Accordingly, the materials having higher temperature tend to lose its heat faster than cold materials. This principle is known as Mpemba Effect in physics. Therefore, if we want to understand which material insulate or conduct heat better, we should look at the percentage of heat loss rather than looking at the differences between initial and final temperatures.)

- Some groups found that metal cup is better insulator than others. How can this observation be explained? (This can be also related with Mpemba Effect. Initial temperature of water in metal cup may be lower than other cups. Therefore, heat loss can be less in metal cup. Likewise, it is possible that water in metal cup lost heat earlier than it was measured by students. Therefore, actual initial temperature of water in metal cup might be higher than what students measured.)

- In this activity, teacher provided one thermometer to each group. The number of thermometer should be equal to number of cups that each groups used. For example; if a group has three cups in their experiments, teacher should give them three thermometers. If teacher gives them one thermometer, they will not measure heat losses of different cups in same time. Therefore, they may not compare the heat insulation of different cups. In conclusion, I will bring more thermometers for this activity in main study.

- In this activity, only some of the students participated in the activity. Some of the students are not voluntarily to engage in argumentation.

- Students do not report what their peers think in their written arguments.

- I used so many cups in this activity, and so measurements got so much time. I can decrease number of cups to be used in main study.

Above mentioned pilot study notes guided the preparation of the second week activities of the main study. More specifically, pilot study showed that students liked the snowman activity as pre-activity and all students actively discussed in this activity. The same activity was decided to be used in main study. One hour was given for snowman activity and experiment design in pilot study. Students were able to design their experiments in given time. Therefore, one hour was again given in main study for snowman activity and design of experiment. In pilot study, conduction of experiment lasted one hour, and whole class discussions lasted one hour too. Due to fact that one hour was sufficient for each of the activities, one hour was given for each activity in main study too. As it was planned and conducted in pilot study, one hour was given for the review of the week in second week of the main study. Due to fact that group members damaged the materials like thermometer in pilot study, greater number of thermometers and other stuffs were used in the main study. By this way, students were able to complete their experiments. In pilot study, students did not make precise measurements. Therefore, students were advised to measure things precisely at the beginning of this activity in main study. How volumes of liquids are measured, how thermometer is used was taught to students at the beginning of the second week in main study. Units like millimeter and Celsius and their symbols were also taught to students. The insufficient content knowledge of researcher regarding the relation between initial temperature and heat transfer was eliminated. Mpemba Effect was learnt and this concept was introduced in main study after students engage in argumentation regarding the comparison of heat insulation of different materials. By this way, Mpemba effect assisted students' understanding about heat transfer. In pilot study, some of the students engaged in argumentation and others did not and this was a problem. In order to engage all students in whole class discussions, teacher was advised to ask more questions. By this way, more students might engage in argumentation. Likewise, silent students were encouraged to propose their ideas in main study. It was also said that all ideas are valuable and deserve to be discussed, so they should be shared and discussed in class.

Third week Notes (Thermal Insulating Products); 18.04.18

- I made adjustment on student and teacher templates because there was no experiment in this week. We replaced experiment with an activity.

- In this activity, thermal insulating products for insulation of buildings were discussed because this was one of the objectives found in curriculum.

- I have prepared evidence sheets and added some visuals regarding thermal insulating materials because students were not familiar with these materials and there was limited information in textbook about these materials.

- In this process, students predicted which material should be used in which part of the house without evidence cards at the beginning. This matching prediction lasted one hour. Then, students re-matched materials and the parts of house for insulation using evidence cards. This also lasted one hour. Students made whole class discussion in third class hour. In the last hour, teacher informed students about the criteria when selecting which thermal insulating products are used in which part of the house. This activity lasted four hours as other activities.

- Whole class discussion was better than small groups' discussions. In whole class discussions, although it was not planned in lesson plans, teacher drew a table on board to match insulating products and corresponding parts of house. Students were expected to match related insulating materials and parts of house on this table. Students actively used this table in their whole class discussions. Therefore, same table can be drawn by the teacher in whole class discussions of the main study.

- The number of students who participated in argumentation increased in this week comparing with previous weeks.

- When students make explanations, they tended to use evidence cards provided them.

- The number of home parts for insulation can be decreased because of limited time. Discussing on all parts of the house and their matching with appropriate thermal insulating materials took too much time.

- Projector did not work properly; therefore, I took hard copy print of power point presentations.

Considering the third weeks' notes of pilot study, some revisions were decided to be made while preparing the third week of the main study: drawing a table that assists matching thermal insulating products and related part of house was added to lesson plan for the main study because this table drawn by teacher on board facilitated whole class discussion in pilot study. Due to fact that time was limited, some parts of house and corresponding thermal insulating product was not discussed in whole class discussion in pilot study. Therefore, two parts of the house which are interior trim and exterior trim were removed from the content for the main study. In other word, content of this week was simplified because of limited time. Finally, projector did not work properly in pilot study, and teacher could not use her presentation in pilot study, therefore, researcher brought hard copy materials including the content found in presentations. By this way, the problem caused by projector was solved.

Forth week Notes (Energy Sources); 25.04.18

- This activity did not include experiment; therefore, I made some adjustment on student template.

- Students reveal their prior knowledge about energy sources and they formed their research questions in first hour.

- Evidence cards were given students in second hour and they worked on this evidence while constructing arguments.

- Group presentations and whole class discussions were done in third hour.

- Sometimes, students cannot focus on lesson and video-camera distracts their attention.

- Higher number of students engage in argumentation; however, it is difficult to control students' misbehavior. They sometimes do not respect their peers. Pro-active rules should be emphasized in main study.

- Classroom management sometimes becomes a problem because students feel themselves totally free in activities.

- While some students never engage in activities, some students are always active participants. On the other hand, some other students have recently started to participate in activities. They participate more over time.

- Their oral argumentation is better than written argumentation reflected in student template.

- The points which were not discussed by students were explained by teacher in fourth hour. This activity lasted four hours (one week) too.

The aforementioned notes of pilot study, guided the preparation of the implementation for the fourth week of the main study: in fourth week of the pilot study, the main problem was the classroom management because it was difficult to control students' misbehaviour (e.g., not respecting others, making noise in class). At the beginning of the main study, some pro-active rules were explained to students in student training. Accordingly, it was told that student centred instruction does not mean that students are totally free, and they can do whatever they want in class. Moreover, it was explained that everyone should respect to others in activities. Similarly, it was told that when there is a chaos in class, quality of their argumentation might decrease, and so they could not learn the topic. Another problem was about students' written argument in pilot study. Accordingly, students rarely completed their SWH reports in pilot study, but it was important for their content knowledge, and quality of arguments. Therefore, more attention was given using SWH in main study. While use of SWH was on students' responsibility in pilot study, students were actively monitored by the teacher in main study regarding the use of SWH reports. By this way, they used SWH throughout the study, and they focused their written arguments more in main study comparing with pilot study.

Fifth week Notes (Electrical Conductivity); 02.05.18.

- The activity lasted four hours. A picture representing Benjamin Franklin's experiment about electricity was used to elicit students' prior knowledge. Students actively engaged and discussed in this activity. This activity can be used in main study. In this activity, students observed that metal key conducts the electricity, but wet rope conducts electricity less. This observation let students to understand that today's topic was about electrical conductivity.

- After that they designed their experiments. Students conducted their experiments in second hour and they constructed their group arguments. Whole class discussion was done in third hour. Teacher summarized the activity in forth hour.

- At the beginning of the study, tape recorder was used in order to reveal small group discussion that happens among group members. However, students do not focus on their experiments when tape recorder is used. They either damage the tape recorder, or they focus on tape rather than small group discussions. In other word, using tape recorder distracted their attention and affected their engagement in argumentation negatively. Therefore, using tape recorder did not work as it was planned, so it might not be used in main study. Only video camera can be used to observe students' engagement in argumentation.

- All groups focused on the classification of substances according to their electrical conductivity. They designed similar experiments and reached similar results.

- Their research questions were not totally the same. Some groups asked which materials do conduct electricity. However, some groups asked which materials do conduct electricity better than others. The second question was more problematic because we had no quantitative measurement to claim one substance has better electrical conductivity than others. The only data we had was bulb brightness which is qualitative and this data was highly subjective for the comparison of different substances' electrical conductivity. To sum up, I need to bring ammeter to class and we should explain the function of ammeter in a circuit in main study because students have not learnt about ammeter. If ammeter is not used, questions like second one cannot be answered and discussed properly.

• Some groups examined whether salty water solution (i.e. solution that include water as solvent and salt as solute) conducts electricity or not. Results of this experiment were very interesting and it also shocked us. We all, including teacher and my research mates already knew that salty water conducts the electricity. However, it does not because bulb did not light when we connected salty water to circuit. On the other hand, water was foaming when we used salty water. This showed that there were some reactions, but no electrical conductivity. Likewise, we observed that colour of water returned to yellow because chlorine gas is one of the product of reaction and it is released in this reaction and gives its colour to the salty water. All these observations that there is no light in bulb, foaming water, and colour change of salty water were interesting and mysterious for students.

• There was no error when sugared water was used in experiment. It did not conduct electricity as it was expected.

• The materials provided to students were salty water, sugared water, batteries (3 volt), bulb (1.5 volt), small and thin connection wires, paper clips, aluminium foil. Students were also free to use and test other materials' conductivity.

• In this week, groups had difficulty in proposing research questions, claims, and evidence.

• We assisted some groups which were not able to conduct experiments.

• After this activity, we discussed how we can better design this activity. We made some adjustments regarding the materials we provided to students. Accordingly, we decreased the size of cup and decreased the amount of water, so percentage of salt increased in solution. We replaced 3 volt batteries with 9 volt batteries, by this way; we provided more electricity to the circuit. We removed small thin connection wires and added more durable and thick wires. We also noticed that the distance between cables are important in salty water. When conductive ends of cables are closer to each other, there is less resistance and bulb brightness increases. As response to these adjustments, salty water conducted electricity and bulb lighted. All these adjustments will be used in main study.

Pilot study also provided information for main study regarding fifth week of the research. First of all, Benjamin Franklin's flying kite experiment was very interesting for students and all students shared their idea and discussed with each other to explain the process such as how electricity reach to Benjamin franklin, and whether wet rope conducts or insulates electricity in pilot study. Because of its effectiveness, this activity remained for main study. Next, small group discussions were aimed to be uncovered by use of tape recorder in pilot study. However, existence of tape recorder distracted students' attention in pilot study. For example, the groups in which tape recorder were used engaged in experimentation and small group discussion less than other groups where tape recorder were not used. Therefore, tape recorder was not used in main study. By this way, small groups' discussions were not uncovered too. On the other hand, whole class discussions were uncovered by use of video cameras in main study. While small group discussions informed about one group's argumentation, whole class discussions informed about all groups' argumentation. Therefore, it is thought that focusing on whole class discussion is more comprehensive and informative comparing with focusing on small group discussions. Another problem of this week was about necessary materials. Some groups wanted to compare the degree of different conductors in terms of electricity in pilot study. However, students had no device (i.e. ammeter) to compare objects' electrical conductivity. Therefore, ammeters were included in main study. Before, students used ammeters, their functions and how they are connected to the circuit was explained to the students in main study. By this way, comparison of conductivity level of different objects became possible. Finally, some adjustments were done regarding the salty water experiment based on the feedbacks obtained in pilot study.

The problem was that salty water did not conduct electricity in experiments done in pilot study; however, it is known that salty water conducts electricity. Therefore, some adjustments were done for the main study regarding this experiment. In line with this, voltage of battery was increased. By this way, more electrical current pass through circuit. Likewise, saturation of salty water solution increased. By this way, more ions were added to experiments that facilitated electrical conduction. Moreover, the distance between ends of cables was decreased. By this way, resistance of the circuit decreased. To sum up, these three adjustments provided bulb brightness in salty water experiments. All these adjustments were used in main study.

Sixth week Notes (Factors affecting light brightness); 09.05.18

• The focus of this week was how characteristics of wire affect the light brightness.

• Research questions were provided to groups. Two groups focused on the effect of wire length on light brightness, other group studied on the effect of cross sectional area (thickness) of wire on light brightness, and other group worked on the effect of different type of wires on light brightness.

• We provided electric circuits to each group and assisted groups when they conducted their experiments. However, groups reached unexpected results. For example, some groups found that when we increase the length of wire, light brightness increases or it does not change. The same inaccurate results were obtained in other groups too. The materials that we provided to students may not work properly. Therefore, it might be better to use simulation programs to show students the effect of wire characteristics on light brightness.

Last week of the pilot study also assisted preparation of the main study. Accordingly, the experiments were not conducted successfully in pilot study. For example, increasing length of the wire does not result with decrease on bulb brightness. Due to fact that, laboratory experiments of this week was replaced with a simulation for the main study. In this simulation, students conducted same experiments in virtual context. All the material in simulation was same with laboratory experiment. While students could not get expected results in their laboratory experiments in pilot study, students participated in main study were able to reach expected results when they did same experiments in virtual context.

## 3.5. Data Collection

Data collection part includes two main sub-parts. In first sub-parts, instruments used in data collection are presented. Then, treatment conducted in main study is explained.

### 3.5.1. Instruments.

Four different instruments which are content knowledge tests, epistemological belief questionnaire, science process skills test, and observations were used in current study. The following part explains each instrument consecutively.

#### *3.5.1.1. Content Knowledge Tests.*

Content knowledge tests included matter and heat unit content knowledge test and electricity unit content knowledge tests. Each test was prepared by the researcher. Each test included 20 multiple-choice test items. Details of content knowledge tests are presented in next part.

##### *3.5.1.1.1. Matter and Heat Content Knowledge Test (MHCKT).*

The MHCKT was developed to assess students' content knowledge in the unit of matter and heat. During the development of the test, firstly, a table of specification was prepared. While preparing the table, attention was given to curricular objectives taking Bloom's Taxonomy of educational objectives into consideration (Bloom, 1956). Then, twenty multiple choice items were developed in line with the table of specification. Test questions were written by researcher considering the content suggested by the curriculum (MONE, 2013). Each item included 4 alternatives. The prepared items were examined by one professor in physics and one professor in science education. Based on the feedbacks provided by the experts regarding accuracy and quality of the items and the appropriateness of the proposed match between objectives and items, necessary revisions were made. Of 20 items developed, 6 were prepared to address the objectives at knowledge level. There was just 1 item at application level.  Majority of the items (8 items) targeted the objectives at Analysis level. While 3 items were at Synthesis level, 2 items were at Evaluation level. In Bloom's taxonomy, knowledge level represents lowest level of thinking. At this level, students are expected to just recall the information.  In the MHCKT, only 6 of the items were at knowledge level, and remaining items were at application level or beyond.  Therefore, it can be asserted that the MHCKT emphasizes higher order thinking skills.

After revising the items, based on the experts' opinions, the MHCKT was pilot tested with 7$^{th}$ grade students who already learnt about the topics related to matter and heat unit. When the pilot study was conducted, 6$^{th}$ grade students did not learn about these topics yet. A total of 156 7$^{th}$ grade students from two schools located in Çankaya district completed the matter and heat content knowledge test in pilot study. Students completed the matter and heat unit content knowledge test in 20 minutes in average

In the pilot study, item test difficulty and item discriminating power were calculated for each test item. Item difficulty shows the percentage of people passing the item. If percentage is high, this shows that item is easy. On the other hand, if number of right answer is low, this shows that item is difficult (Miller, Linn, & Gronlund, 2009). Although, the computed values are used as an indication of item difficulty, the real difficulty is linked to the effectiveness of instruction. Thus, a higher value can simply indicate that the instruction was effective (Popham, 2005). Table 3.3 shows items' difficulty levels, item discriminating power, and Bloom Taxonomy level that items corresponds.

Table 3. 3

*Matter and Heat Content Knowledge Test Item Difficulties, Item Discriminating Power and Corresponding Bloom Taxonomy Levels*

| Matter and Heat Test Item Difficulty | | Item Discriminating Power | Bloom Taxonomy Level |
|---|---|---|---|
| Item 1 | 0.66 | 0.46 | Knowledge |
| Item 2 | 0.66 | 0.67 | Analysis |
| Item 3 | 0.53 | 0.69 | Analysis |
| Item 4 | 0.22 | 0.12 | Analysis |
| Item 5 | 0.57 | 0.72 | Knowledge |
| Item 6 | 0.35 | 0.49 | Analysis |
| Item 7 | 0.30 | 0.41 | Synthesis |
| Item 8 | 0.31 | 0.05 | Knowledge |
| Item 9 | 0.40 | 0.36 | Evaluation |
| Item 10 | 0.38 | 0.51 | Analysis |
| Item 11 | 0.28 | 0.44 | Synthesis |
| Item 12 | 0.30 | 0.64 | Knowledge |
| Item 13 | 0.50 | 0.77 | Knowledge |
| Item 14 | 0.40 | 0.62 | Analysis |
| Item 15 | 0.36 | 0.44 | Analysis |
| Item 16 | 0.31 | 0.44 | Synthesis |
| Item 17 | 0.43 | 0.62 | Knowledge |
| Item 18 | 0.36 | 0.56 | Application |
| Item 19 | 0.38 | 0.66 | Analysis |
| Item 20 | 0.25 | 0.36 | Evaluation |

According to this table, the items 4, 11 and 20 are more difficult than other items. Item difficulties ranged from .22 to .28. While item difficulty considers whether an item is easy or difficult, item discriminating power shows how an item discriminate high achievers from low achievers. In discriminating power calculation, upper %25 percent of students (n=39) and lower %25 percent of students (n=39) were used. The following formula was used to calculate item discriminating power of each item: D= (RU-RL)/ (0.5T) where RU is the total number of right answers of the upper group, RL is the total number of right answer of the lower group, and T is the total number of students trying the item.  According to Ebel and Frisbie (1991), while D value between .30 and .39 suggest reasonably good item, the D value greater than or equal to .40 indicates very good item. Based on these criteria, Table 3.3 shows that, all items except for item 4 are reasonably good or very good items. Although, the item 4 had low discriminating power, since the discriminating power is not an indication of item validity and a defective item (Miller et al., 2009), it was decided to be retained in the test. In fact, Miller et al suggested that, if there are no technical defects in the item and if it assesses an important objective, then it should not be removed from the test.

In addition, according to table 3.3 all the items had positive discrimination that means the items were responded correctly more often by the students performing well on the overall test than the students who performed poorly (Popham, 2005).

In order to determine the reliability, Kuder-Richardson Formula 20 (KR-20) was applied which provides a measure of internal consistency. KR-20 is used when items have different degree of difficulty and items are dichotomously scored. In the pilot study, KR-20 value was found as .70 suggesting sufficient internal consistency. In the main study, KR-20 value for pre-test was .64 and KR-20 for post-test was .76. Overall, the matter and heat unit content knowledge test appeared to possess good psychometric properties with sufficient validity and reliability evidences. A copy of the test was presented in Appendix E.

##### *3.5.1.1.2. Electricity Content Knowledge Test (ECKT).*

The ECKT was developed to assess students' content knowledge in the unit of Electricity. During the development of the test, a table of specification was prepared considering curricular objectives in relation to Bloom's Taxonomy of educational objectives. While some of the curricular objectives were directly related to conceptual understanding, others were related to science process skills. In line with the table of specification, a 20 multiple choice item test was constructed. Researcher prepared these questions and content suggested by curriculum (MONE, 2013) was considered when questions were written. Each item had four alternatives. The items were reviewed by one professor in physics and one professor in science education.

Based on the experts' feedbacks regarding accuracy and quality of the items and the appropriateness of the proposed match between objectives and items, necessary revisions were made and the test was pilot tested. Of 20 items pilot tested, 11 were prepared to assess students' conceptual understandings and 9 were prepared to assess students' science process skills (see Table 3.4). In addition, as shown in the table, 4 of the items were at knowledge level and remaining items were all at comprehension level (2 items) and beyond (14 items). Thus, because the majority of the items targeted complex learning outcomes, the test can be considered as assessing higher order thinking skills.

The pilot study was conducted with a total of 195 $7^{th}$ grade students who learnt about electricity topic in the previous year. The students were from a middle school located in Çankaya district. Students completed the electricity unit content knowledge test in 20 minutes in average. The findings in the pilot study, regarding item difficulty and item discriminating power were presented in Table 3.4. According to the table item 12, item 14 and item 20 were relatively more difficult than other items. In general, it was thought that questions were in optimum level of difficulty and none of the questions were removed from test. Table 3.4 also shows item discrimination index of electricity content knowledge test. Upper %25 students (n=49) and lower %25 students (n=49) test results were used to calculate item discrimination index. All the items , except for item 12,14, and 20, were found to have reasonably good or very good items. The three items which can be considered as marginal (Ebel & Frisbie, 1991) were decided to be retained in the test, because the discriminating power is not an indication of item validity, and a defective item (Miller et al., 2009). Indeed, according to Miller et al, if there are no technical defects in the item and if it assesses an important objective, then it should not be removed from the test.

Table 3. 4

*Electricity Content Knowledge Test Item Difficulties, Item Discriminating Power Index, Item Contents and Corresponding Bloom Taxonomy Levels*

| Electricity Test Item | Difficulty | Item Discriminating Power Index | Item Content | Bloom Taxonomy Level |
|---|---|---|---|---|
| Item 1 | 0.36 | 0.41 | Conceptual Understanding | Knowledge |
| Item 2 | 0.55 | 0.82 | Conceptual Understanding | Knowledge |
| Item 3 | 0.62 | 0.78 | Conceptual Understanding | Comprehension |
| Item 4 | 0.68 | 0.76 | Conceptual Understanding | Application |
| Item 5 | 0.65 | 0.69 | Conceptual Understanding | Application |
| Item 6 | 0.70 | 0.69 | SPS-Classify | Analysis |
| Item 7 | 0.60 | 0.80 | SPS-Research Question | Analysis |
| Item 8 | 0.34 | 0.43 | SPS-Identify variables | Analysis |
| Item 9 | 0.55 | 0.82 | Conceptual Understanding | Synthesis |
| Item 10 | 0.44 | 0.69 | Conceptual Understanding | Synthesis |
| Item 11 | 0.32 | 0.55 | SPS-Designing Experiment | Evaluation |
| Item 12 | 0.25 | 0.24 | Conceptual Understanding | Knowledge |
| Item 13 | 0.48 | 0.72 | Conceptual Understanding | Comprehension |
| Item 14 | 0.24 | 0.28 | Conceptual Understanding | Application |
| Item 15 | 0.38 | 0.59 | SPS-Identifying variables | Analysis |
| Item 16 | 0.30 | 0.49 | SPS-Identifying variables | Analysis |
| Item 17 | 0.47 | 0.73 | SPS-Formulating hypothesis | Synthesis |
| Item 18 | 0.36 | 0.45 | SPS-Designing experiment | Synthesis |
| Item 19 | 0.36 | 0.39 | SPS-Formulating hypothesis | Evaluation |
| Item 20 | 0.27 | 0.22 | Conceptual Understanding | Knowledge |

In addition, as a measure of reliability, Kuder-Richardson Formula (KR-20) was used because the items were dichotomously scored and they were of different difficulty. KR-20 value for the electricity unit content knowledge test was found as .72 in the pilot study revealing a sufficient internal consistency. Likewise, KR-20 values for the main study were found as .77 for both in pre-test and post-test. A copy of the electricity unit content knowledge test was presented in Appendix F.

#### *3.5.1.2. Epistemological Belief Questionnaire (EBQ).*

The EBQ developed by Conley et al. (2004) was used to assess students' epistemological beliefs. It includes 26 items on five-point scale in four dimensions, namely source (n= 5, e.g. "If you read something in a science book, you can be sure it's true.", α = .81), justification (n =9, "A good way to know if something is true is to do an experiment." α = .65), development (n= 6, e.g. "New discoveries can change what scientists think is true.", α = .57, and certainty (n= 6, , e.g. "All questions in science have one right answer.", α = .78).

The EBQ was translated and adapted to Turkish by Özkan (2008). During its validation for Grade 7 Turkish middle school students, exploratory factor analysis (EFA) was conducted. Results suggested three factors explaining 45.5 % of total variance. Accordingly, although source and certainty belongs to different aspects of EBQ (e.g., nature of knowledge) in the original version, in Turkish version of the instrument it was decided to combine these two factors and consider them as a single factor. In addition, two items, which were found not to contribute well to the total variance reducing the reliability, were removed from the questionnaire. The reliability for the overall questionnaire was found to be .76. Confirmatory factor analysis (CFA) also supported the three factor structure (GFI=0.92, AGFI=0.91, RMSEA=0.06, S-RMR=0.06).

In the current study, 24 item Turkish version of the EBQ (Özkan, 2008) was used. (see Appendix G). Regarding reliability, Cronbach alpha value for whole scale was found as 0.81 for pre-test and 0.91 for post-test. Cronbach alpha value for combination of source and certainty dimensions were found as 0.75 in pre-tests, and 0.87 in post-test. Development dimension's Cronbach alpha value was 0.60 in pre-test and 0.81 in post-test. Finally, justification dimension's Cronbach alpha value was calculated as 0.73 in pre-test and 0.89 in post-test. Thus, reliability coefficients were found to be high enough to conduct further analyses.

#### *3.5.1.3. Test of Integrated Process Skills 2 (TIPSII).*

The TIPSII developed by Burn et al., (1985) was used to assess students' science process skills. This test aims to assess middle school and high school students' integrated process skills including identifying variables (12 items), operational definition (6 items), stating hypothesis (9 items), interpreting graphs (6 items), and designing investigation (3 items). Overall, the test included 36 multiple choice items with four alternatives. Content knowledge was not emphasized when TIPSII was prepared. The emphasis was, on the other hand, on the use of integrated process skills. Therefore, participants do not have to be content expert when solving TIPSII questions. During its development, six science educators examined the test in terms of content validity. In this process, educators separately answered the multiple choice questions, matched the questions and objectives and made comments for modification of the test. A total of 459 students from Grade 7 to Grade 12 completed the test. It took 40-50 minutes for middle school students to complete the test. Cronbach alpha level was found as 0.86. Results of the study also showed that TIPSII scores significantly correlated with TIPSI which was prepared by same researchers to assess students' science process skills. Accordingly, both TIPSI and TIPSII were prepared to assess integrated process skills; however, content of the questions were different from each other for these two tests. This high correlation between test scores supported the construct validity of TIPSII.

The TIPSII was translated and adapted to Turkish high school students by Geban, Askar, and Özkan (1992). Number of item was 36 as in original scale and five dimensions of the test which are identifying variables, stating hypothesis, operationally defining, designing investigation, and interpreting graphs also existed in adapted version. Reliability coefficient was calculated as 0.81 for the whole test. Then Can (2008) applied this scale including 36 items to 227 middle school students and removed 10 items from the original scales because of their low item discriminating power. Final version of the adapted TIPSII had 26 items and the researcher found KR-20 for the whole instrument as 0.80. In this final version, defining operationally includes 3 items, identifying variables includes 11 items, stating hypothesis includes 6 items, interpreting graph consists of 3 items and designing experiment includes 3 items. the In the current study, short version of the TIPSII validated for middle school students by Can (2008) was used (see Appendix H). KR-20 value was found to be .83 in pre-test and .86 in post-test.

Actually, TIPSII data should have been examined considering science process skills' dimensions (e.g., defining operationally). However, reliability scores of dimensions were too low because of few numbers of items assessing corresponding dimension. Therefore, related analyses were not performed considering dimensions of TIPSII. On the other hand, analysis was conducted based on overall scores of TIPSII. The same issue was observed in previous research regarding low reliability scores of TIPSII dimensions and researchers conducted analysis based on total scores (Gök, 2014). This situation is a limitation of the study and interferes with the determination of the impact of the ABI in terms of the dimensions of the science process skills.

#### *3.5.1.4. Observations.*

Up to this point, instruments used to collect quantitative data were presented. On the other hand, observations were used to collect qualitative data in this study. The data obtained from observations were mainly used to answer forth and last research questions that examine $6^{th}$ grade students' engagement in scientific argumentation and argumentation schemes.

Observations are the first hand data that is collected in natural environment (Merriam, 2009). Video camera records were used in this study to collect observation data both in pilot study and main study. Ethical permissions were obtained from university, Ministry of National Education, and school administration for both pilot and main study to record class instances with camera before study started. Moreover, teacher contacted with students' families and all parents and the students voluntarily accepted use of video cameras. Accordingly, each of the four classes was video-recorded four lesson hours (each lesson is 40 minutes) in a week and study lasted six weeks for both pilot and main study. Same observation process was followed both in pilot and main study. Therefore, each class was observed 24 lesson hours which is equal to 960 minutes for each class. Video cameras were placed at the corner of the laboratory, therefore it did not get students' attention too much, and thus it did not affect students' behavior negatively. Data was collected on April and May 2018 for pilot study and it lasted six weeks. Similarly, data collection process via video-cameras started in at the beginning of March 2019 and finished six weeks later in April 2019 for the main study.

In main study, two of the classes were selected randomly and these classes' video records were used for data analysis. Accordingly, each lesson's video records were watched by researcher. The discussion episodes were identified in each lesson. Identification of discussion episodes lasted one month and this process was done in June 2019. Then, researcher transcribed each discussion episodes one by one. Transcription of data lasted two months. This process started in June and finished at the end of August 2019.

Researcher and research mates were also in classes when ABI instruction was implemented. Researcher position was being a complete observer in this study. According to Merriam (2009) complete observer means that researchers do not participate in any activity and they just observe what is happening in class. Researcher and research mates sat at the corner and observed teacher's implementation of ABI and students' whole class discussions. When students needed help in their experimentation process, research mates provided feedbacks to the students. Likewise, research mates assisted distribution of materials like experiment equipment and evidence cards. Existence of researcher and research mates did not affect students' participation because students got familiarity with researcher in student training conducted prior to main study. Moreover, teacher was not affected by researcher and camera records because teacher got familiarity to this research from pilot study and conversations held with researcher in teacher training. Data collection tools and corresponding research questions were summarized in table 3.5. Details of the treatment done in main study are explained in next part.

Table 3. 5
*Summary of Data Collection Tools and Corresponding Research Questions*

| Research Questions | Data Collection Tools |
|---|---|
| 1. Is there a change in $6^{th}$ grade students' content knowledge from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)? | Heat and Matter Content Knowledge Test<br>Electricity Content Knowledge Test |
| 2. Is there a change in $6^{th}$ grade students' epistemological beliefs from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)? | Epistemological Beliefs Questionnaire (Conley et al., 2004) |
| 3. Is there a change in $6^{th}$ grade students' science process skills from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)? | Science Process Skills Test (Burn et al., 1985) |
| 4. What is the nature of $6^{th}$ grade students' arguments when they are analyzed based on argumentation schemes? | Observations |
| 5. What is the nature of $6^{th}$ grade students' engagement in argumentation process in ABI treatment? | Observations |

### 3.5.2. Treatment.

Teacher templates and feedbacks obtained from pilot study determined how the treatment to be carried out in the main study. Accordingly, phases of the teacher template were considered while implementing the ABI treatment. Each week's activities are presented based on the phases of teacher template (e.g., pre-lab activities) in this part.

Prior to main study, researcher met with students and there has been a conversation between researcher, teacher and students. Content of this conversation was scientific debate, inquiry, importance of different ideas in science, scientists, similarities between students and scientists, how scientists work, and science process skills. Students voluntarily participated in this conversation. After that researcher informed students about argument-based inquiry and student templates. By this way, students were able to understand what are expected from them in this study. Next, proactive rules were shared with students. Accordingly, students were supposed to respect each other, attack others' ideas (not their identities), follow the evidence, and avoid personal bias. At the end of conversation, researcher brought a bell jar that was full of water and there was a predict-observe-explain activity about density topic. Researcher asked students to predict which object would float or sink. Students shared their predictions and reasonings with class. Then, researcher threw different objects in water, and students observed their position relative to water. After that students explained the reasons of their observations. Some ideas were opposed to each other and students engaged in argumentation in this conversation. Using this activity, researcher explained that this was an inquiry activity and students constructed and evaluated their arguments in argumentation process. Researcher also provided information regarding the components of argument (e.g., claim, counterclaim, and data) using students' ideas. This conversation, held prior to main study, lasted four lesson hours (one week) in each class involved in the study.

In the first week of the treatment, the core idea was the classification of matters based on their heat conductivity. First phase in the implementation was eliciting students' prior knowledge. Accordingly, teacher asked students which type of cooker should be used when we go to picnic to cook foods. The alternatives provided were clay pot and metal pot. There was a class discussion to answer this question. Students proposed different claims and their justifications. They used counterarguments for the claims which they did not agree. Teacher started another discussion by asking which kind of fork (e.g., metal, plastic, wood) should be used if we want to mix the cooked food. Another argumentation process started and students tried to refute opponent ideas. By this way, teacher was able to elicit students' prior ideas regarding heat conductivity. In second phase (i.e., pre-lab activities), teacher asked students to construct a research question they would like to test. After each group formed their research question, these research questions were analysed in terms of appropriateness. Both teacher and other groups shared their ideas about the research questions. The research questions starting with how and why questions were eliminated and groups set some other research questions. For example; students could not answer the research question "Why does metal conduct heat better than wood?" by doing experiment, so this question was not suitable for research. Likewise, research questions which were not related with curricular objective (i.e., classification of substances based on heat conductivity) were also abandoned. In line with this, students formulated research questions such as “which type of spoon does conduct heat better?” Then groups formed their own experimental design to answer their research questions. They stated their hypothesis, independent, dependent, and control variables and research designs. In third phase (i.e., laboratory activities), students performed their experiments. The materials which are easily accessible in school were chosen and provided to the students. These materials were metal spoon, wood spoon, glass spoon, plastic spoon, beaker, aluminium foil, thermometers, clock, solid oil, and hot water. Students did not have to use all these materials and they used the materials which were consistent with their experiments. Students did not directly use hot water; teacher provided hot water and poured it in beakers because of safety rules. However, groups determined the amount of hot water to be used. Although groups’ research questions were similar, their data collection processes were different from each other. For example, some groups put different spoons in hot water; they added solid oil on spoons and observed which spoon’s solid oil melt faster than others. They concluded that faster melting of solid oil is result of better heat conduction. On the other hand, some other groups put spoons in hot water, they waited in few minutes. Then, they removed spoons and put them on table. After that they added the solid oil on different spoons and they decided which spoon conducts heat better than others. In conclusion, students were able to design and conduct their experiment in general. However, some groups could not able to design and conduct their experiments. Teacher assisted these groups in designing and conducting experiments. Students reported their research questions, design of the experiments, hypothesis, and variables in student template. Likewise, students recorded their observations such as temperature change of water over time, melting of solid oil to student template. After experiments were conducted and observations were reported in students’ templates, phase-4 (negotiation phase-1) started. In this phase, students individually constructed their arguments reporting their claims and evidence. For example, some students claimed that metal is better heat conductor because solid oil on metal melted earlier than other oils. In phase-5 (negotiation phase-2), students compared their individual arguments with group members and groups formed their group arguments based on group discussions. Students mainly had similar observations and their individual arguments were similar to each other.

When group members did not discuss further, teacher guided these groups and asked questions such as “What could be the reasons of these observations?”, “Did you consider all the observations?”, “Did you falsify or retain your hypothesis?” In phase-6 (negotiation phase-3), groups presented their experiments, findings and arguments to other groups in order. There were whole class discussions. Other groups’ examined and evaluated presenter group’s arguments. For example, some groups found that plastic spoon was better heat conductor; on the other hand, some others found that metal spoon was better heat conductors. Groups examined the evidence and data collection procedures further to reach a consensus. After that groups focused on theoretical explanation regarding why some substances conduct heat better than others. Teacher engaged in these theoretical explanations and students summarized that some substances are better heat conductors because their particles have an order and these particles are closer to each other that facilitate heat transfer. Then, teacher said that substances can be classified based on heat conductivity. The substances that transfer heat better are heat conductors; on the other hand, substances that do not conduct heat well are heat insulators. In phase-7 (negotiation phase-4), students wrote their reflections regarding what they learnt in this activity.

While heat conductors were emphasized in the first week, the core idea of the second week was heat insulators. In the first phase, students’ prior knowledge regarding heat insulation was elicited through an activity about snowman. Teacher showed students two different pictures. In first picture, there was a snowman without any clothes and there was a snowman wearing coat in second picture. Related with these pictures, teacher also added that the weather is sunny and outer environment’s temperature is higher than snowman’s temperature. Teacher asked students which snowman melt earlier or faster. By asking this question, teacher aimed to understand students’ prior knowledge about heat insulations. Most of the students claimed that snowman wearing coat melt faster because the air is hot and hot air gets snowman melts, moreover, students added that coat make snowman hotter causing faster melting. On the other hand, few students used correct argument claiming that coat that snowman block heat transfer from outer environment to the snowman, therefore, snowman wearing coat melt slower than snowman without coat. After this class discussion, teacher asked students to design experiments about heat insulation as a group in second phase (pre-lab activities). Teacher provided students with materials. These materials were easily accessible materials which were foam cup, plastic cup, glass cup, metal cup, thermometers, and hot waters. Students formed their research questions such as which of the materials is the best heat insulator and what is the relationship between cup type and heat insulation. Then, groups designed their experiments and shared their designs with the rest of the class. More specifically, each group told their research questions, variables, hypothesis, and data collection procedure. Other groups and teacher provided feedbacks about experiments and groups made adjustments on their experiments based on feedbacks. In phase 3, all groups conducted their experiments. As in the first week, teacher poured hot water to different cups for safety. Students recorded temperature differences on thermometer over time to student template as their observations. Moreover, students recorded their qualitative observations such as “when I grabbed the metal cup, it burnt my hands.” In phase 4 (negotiation phase-1), students individually report their claims and arguments. For example, one student wrote that I hypothesized that foam cup is the best insulator; however, my observations showed that heat loss in plastic cup was less than foam cup. Therefore, my claim is plastic cup is better heat insulator. In phase 5 (negotiation phase-2), students discuss their individual arguments with group members and they formed group arguments. While groups formed their group arguments, there was no so much discussion. Teacher moved around the groups and asked them whether their hypothesis was falsified. Moreover, teacher asked students other possible factors affecting their experiments. Such questions provoked students and encouraged them to think and discuss further. After group arguments were formed, whole class discussions started in phase 6 (negotiation phase-3). Groups presented their arguments, experiments and findings in order. Other groups criticized or supported the presenter groups. When groups had same research question and reached different claims, discussions were at the top level. For example, students used temperature difference to decide which cup is better heat insulator; however, temperature difference was the lowest in foam cup in one group and it was lowest in plastic cup in another group. So, groups could not reach a consensus. Then, one of the students claimed that they should focus on percentage of heat loss for each cup (i.e., Mpemba Effect) rather than focusing on only heat loss. Then, class applied this suggestion, and they reached the conclusion that percentage of heat loss in foam cup is more than other cups for each group. Therefore, all groups reached a consensus that foam cup is better heat insulator and foam can be used for insulation. Moreover, whole class discussions included the factors affecting heat loss of the cups such as initial temperature of the hot water, size of cup, ambient temperature and so on. After that in phase 7 (negotiation phase-4), teacher summarized the lesson and students wrote their learning as reflections to student templates. Then, teacher re-asked the snowman question asked at the beginning. Most students answered that snowman wearing coat melts later than snowman wearing nothing because it blocks heat transfer.

Third week's core idea was thermal insulating products used in different parts of the house. No experiment was conducted in this week and evidence sheets were given to students. In phase 1 (eliciting prior knowledge), teacher showed a picture of a house to students. In this picture, students were asked to compare heat loss of different parts of the house. After showing different parts of the house (internal wall, external wall, ceiling, installation pipe and solar panel), teacher asked students which thermal insulating products (plastic foam, tar, glass wool, silicone wool, rock wool, and wood wool) should be used in which part of the house. Students were expected to match thermal insulating products and corresponding parts of the house using their background knowledge and predictions in this pre-activity (phase 2). Due to fact that students did not have prior knowledge about properties of thermal insulating products and the areas that they are used, students either mismatched these products and corresponding house parts or they could not match them. In phase 3 (main activity), teacher presented the evidence cards to each groups. No research question was assigned to groups. All groups were free to produce their own research questions. Group research questions were like "Which thermal insulating products should be used for internal wall?", "In which part of the house should stone wall be used?" After groups formed their research questions, they started to gather data by analysing evidence cards. Evidence cards were prepared based on textbook suggested by curriculum (see Appendix I). The information found in evidence card included four different criteria for the selection of thermal insulating products. These criteria were flammability, area of usage, durability and economic cost. While addressing their research questions, students were asked to consider all these four selection criteria. In phase 4 (negotiation phase-1), students individually wrote their claim and evidence in student template. For example, one student wrote that glass wool should be used in internal wall because it does not burn easily; it can be used for many years, it is economical and one of the house part that it can be used is internal wall. In phase 5 (negotiation phase-2); group members discussed their claims to construct group argument. For example, individuals criticized each other when they selected different thermal insulating product for same part of the house. While students criticized each other, they considered selection criteria given in evidence cards. After groups formed group arguments, they engaged in whole class discussions in order (phase 6, negotiation phase-3). Again, different groups considered selection criteria for thermal insulating products when supporting, defending and criticizing their arguments. Next, teacher summarized the lesson pointing out thermal insulating products and their selection criteria for heat insulation. Students write their learning as reflection in student template (phase 7, negotiation phase-4).

Forth week's core idea was sources of energy. As in third week, there was no experiment in forth week and researcher prepared evidence cards for students to make them construct their arguments. The same research questions were formed by all groups in this week. The research questions were "Which energy source should we use? and "Which energy source should not we use? " Sources of energy was obtained from textbook which are wood, coal, petroleum, natural gas, solar energy, wind energy, geothermal energy, and biomass. Food, hydroelectric energy and nuclear energy were excluded as sources of energy because they were not included in curriculum. At the beginning of the lesson, teacher showed three pictures to the students. There was a car going on road in first picture, a bulb lightning in a room in second picture and a man running on the street in third picture. Teacher asked students to discuss common points of these three pictures. Students were expected to explain that all objects in the pictures need energy to sustain their situation. After students connected the topic into energy, teacher asked the sources of energy used in these pictures. By this way, students' prior knowledge about sources of energy was elicited (phase 1, eliciting prior knowledge). Then, teacher presented different sources of energy and asked students which energy sources we should use and which energy sources we should not use. Using their background knowledge, students tried to answer teacher's question without any evidence cards. For example, one of the students claimed that wind energy should be used because it is free, and petroleum should not be used because it contaminates the sea, river and ocean. Similarly, other students proposed their initial ideas regarding which energy sources should or should not be used. They also discussed with each other prior to the main activity about these energy sources even though no evidence cards were given them. This activity was used as pre-activity (phase-2). Then, teacher distributed evidence cards to each groups. Evidence cards were prepared by the researcher and included definitions, advantages and disadvantages of each energy sources. Evidence cards used in week 4 is presented in Appendix D. Students analyzed each evidence cards and so they collected data to answer the research question (Which energy source should we use, which energy source should not we use?). This data collection process using evidence cards was main activity (phase-3). After that, students individually wrote their claims and evidence to student template (phase-4, negotiation phase-1). For example, one of the students wrote that we should use solar energy because solar energy is free, we can benefit from solar energy to heat our houses, heat water and cook food. Moreover, we can transform solar energy to other energy types. On the other hand, we should not use wood as sources of energy because it is not efficient to use. Seventy five percent of burned wood turns into ash and fog before it is used. In phase-5 (negotiation-2), individuals discussed with their peers to form group arguments. For example, one of the group members thought that they should not use solar energy because sometimes there is no sun and we cannot benefit from sun. Likewise, some places do not get enough solar energy to use it; therefore solar energy is not usable for these places. Groups constructed their arguments depending on these feedbacks and alternative ideas. After group arguments were formed, whole class discussion started (phase-6, negotiation phase-3). Groups presented their group arguments in turn. For example; there was a debate between groups in class-1 regarding whether we should use wood as source of energy or not. Accordingly, one of the groups claimed that they did not want to use wood as energy source justifying that burning wood causes climate change. On the other hand, another group members rejected and added that “Not only wood, but also petroleum and natural gas cause climate change, why did not you select petroleum and natural gas, if your focus is climate change?” and discussion continued. After debates ended, teacher wanted students to read energy sources part from textbooks. Teacher asked students whether there was any new term that they were not familiar with. Students mentioned renewable and non-renewable energy sources. Class engaged in further discussion about classification of energy sources as renewable and non-renewable energy sources. In last phase (negotiation phase-4), students wrote what they learnt in this week as their reflections in students template. Energy sources were the last topic of Matter and Heat unit. After this week, electricity unit started and lasted two weeks.

Core idea of fifth week was electrical conductivity; students were expected to classify substances as conductors and insulators. At the beginning of the lesson, teacher showed a picture representing Benjamin Franklin’s electricity experiment which was done to understand whether lightning is electrical energy or heat energy. In this picture, Benjamin Franklin grabs a rope in a rainy day, and a flying kite is tied to this rope. Besides, there was a key on the rope and this key shines. According to this initial activity, students also decided which objects in the picture are conductor and which objects are insulators in terms of electricity. In this point, students make inferences based on the data found in picture (e.g., key shines). Likewise, students use their background knowledge when they explain electrical conductivity of objects. For example, students might think that key is conductor because it shines; on the other hand, there can be conflict in class whether wet rope is conductor or insulator. By this way, students’ prior knowledge about electrical conductivity was elicited (phase 1). In these early discussions, students discussed whether lightning is electrical or heat energy, whether key is a conductor, whether rope is a conductor, whether electricity reaches kite, if it reaches kite how it is possible, whether electricity reach Benjamin Franklin how he was affected by electricity. Many of the students voluntarily engaged in this early discussion. After that, teacher explained what happens in picture and made a connection with the topic of electrical conductivity. Then, teacher provided materials to students and asked students to write their research questions, experimental designs, variables and hypothesis as pre-lab activities (phase 2). The materials provided to groups were battery (9v), bulb (1.5v), socket, connecting cable, paper clip, ceramic cup, glass cup, duster, aluminium foil, metal spoon, plastic spoon, wood spoon, three beakers, salt, sugar, and water. Students were free to write their research questions but there were three conditions that they had to obey. First, research question should be relevant to electrical conductivity. Second, research question should be answered by use of provided materials which means students should not need any additional material to conduct their experiments. Third, questions should not start with how and why words that requires theoretical explanations which cannot be derived from direct observations or experiments. Next, students prepared their research questions such as which materials do conduct electricity? Which of the material does conduct electricity better than others? Which materials are conductors? Which materials are insulators? Next, students designed their experimental set up and conducted their experiments (phase-3, laboratory activities). They recorded their data individually as their observations in student template. Then, they wrote their claims and evidence, this assisted them to construct individual arguments (phase-4, negotiation phase 1). For example, one of the students wrote that sugared water and tap water do not conduct electricity because bulb did not light when these materials were connected to electrical circuit. Then, small group discussions started and group members constructed group arguments (phase 5, negotiation phase 2). When groups did not discuss and they all had consensus, teacher interacted with groups and asked further discussion questions such as what could be reason causing sugared water not to conduct electricity? These questions assisted further discussions among group members. After group arguments formed, groups presented their arguments in order. By this way, whole class discussions started (phase 6, negotiation phase 3). When students tried same materials, but got different results, there were firm discussions. For example, one group found that salty water conducts electricity; on the other hand, another group found that it does not. Moreover, these groups were knowledgeable that salty water conducts electricity because this knowledge was available in textbook. Although all groups knew it conducts electricity, groups sustained supporting their experiments and findings. Then, students discussed possible causes of different results. In fact, some of the students reached theoretical explanation for the conditions that salty water conducts electricity. For examples, students claimed that there is a chemical reaction in salty water because water foamed and it turned to yellow colour when cables put in water and this chemical reaction provided electrical conduction. Likewise, students claimed that when cable ends close each other, less resistance is applied to the electrical current. Therefore, current passes in circuit and bulb lights. In other word, student constructed the knowledge through discussion. After this whole class discussion, teacher summarized the lesson pointing out Benjamin Franklin case, classification of materials as conductors and insulators, and electrical conductivity. After all, students wrote their learning as reflection in student template (phase 7, negotiation phase 4).

Core idea of the final week of the treatment was the factors affecting bulb brightness. Students learnt the factors such as number of bulbs, number of battery affect bulb brightness in previous year ($5^{th}$ grade level). Therefore, the content was limited with the wire properties which are length of wire, cross-sectional area of wire (wire thickness), and type of wire as factors affecting the bulb brightness. At the beginning of the lesson, teacher asked whether we can change the bulb brightness in an electric circuit. Students discussed the factors affecting light brightness based on their background knowledge learnt in previous year. By this way, students' prior knowledge about the factors affecting bulb brightness was elicited (phase 1). Then, teacher asked whether properties of wire that connect different components of circuit to each other affect bulb brightness. Students tried to answer this question too. Next, teacher wrote three research questions to be tested. These questions were "What is the relation between length of wire and bulb brightness?", "What is the relation between thickness of wire and bulb brightness? and "What is the relation between type of wire and bulb brightness?". These questions were randomly assigned each group. Groups formulated their hypothesis, identified variables and their research design to address their research questions (phase-2, pre-lab activities). After all groups shared their research design and they got feedbacks about research designs from their peers and teacher, they were taken to the computer laboratory because experiments in science laboratory did not work properly in previous year (pilot study). In other word, groups were asked to conduct their experiment in a simulation showing how wire properties affect bulb brightness. Each group worked in different computer to answer their research questions. Teacher, researcher and research mates assisted students when they needed help about use of simulation such as setting electrical circuit. When students collected data via simulation, they recorded their findings to student template (phase 3, laboratory activities). After students collected data, they returned to science laboratory. They individually wrote their claim and evidence (phase 4, negotiation phase 1). For example, a student wrote type of wire affects light brightness because light brightness increased when they changed the kind of wire from iron to copper. Next, groups formed their group arguments by comparing their individual arguments (phase 5, negotiation phase 2). When they reached consensus, teacher asked the group some theoretical questions for further discussions. These theoretical questions were like “Why does copper wire conduct electricity better than iron wire?” After groups constructed their group arguments, they presented their experiments and arguments in turn (phase 6, negotiation phase-3). For example, one of the groups told that when they increased the length of the wire, the resistance decreases because the value on ammeter decreases. This group had misunderstanding that ammeter value shows resistance. Other groups rejected and claimed that if length of wire increases, the resistance increases too and less electrical current pass through wire and the value read on ammeter decreases. After whole class discussion, teacher summarizes the lesson telling students that there is a reverse proportion between length of wire and bulb brightness, direct proportion between thickness of wire and bulb brightness and bulb brightness depends on the type of wire. Then, teacher connected topic to resistance of wire and explained how resistance of wire changes when its properties change. Teacher used some analogies to show the relation between wire resistance and wire properties. For example, teacher asked students to think two different cups having full of water. While first cup has a small hole at the bottom, second cup has a big hole at the bottom. Teacher asked students which cup losses water faster. The expected answer was the cup having big hole losses water faster. Then teacher compared cups having full of water to battery, flowing water to electrical current, and the size of hole to the thickness of wire. When hole size increases, water passes through hole easily. In the same manner, when thickness of wire increases, current passes more and bulb brightness increases. Then, students wrote their learning to student template (phase 7, negotiation phase-4).

### 3.6. Treatment Verification

Verification of independent treatment shows to what extent argument-based inquiry was applied in current research. In line with this, how previous research verified their ABI treatment was investigated. In the relevant literature, it was found that Reformed Teaching Observation Protocol (RTOP) developed by Sawada, Piburn, Falconer, Turley, Benford, and Bloom (2000) was used as a guideline by ABI researchers to verify their treatment (Martin & Hand, 2009). For example, Martin and Hand (2009) used RTOP to assess teachers' implementation of ABI. Researchers claimed that there was consistency between RTOP's 25 items and ABI. Especially, 13 items were directly related with ABI (Martin & Hand, 2009). Accordingly, in the present study, throughout the data collection process, the researcher decided to use the RTOP scale as checklist while observing teacher's ABI implementation for each week. In these observations, researchers focused on 13 key items (e.g. "Student questions and comments often determined the focus and direction of classroom discourse.") which are considered to be essential for the ABI implementation. After each observation, the researcher marked RTOP checklist and shared it with the teacher. The teacher adjusted her teaching considering RTOP feedbacks in the following weeks. By this way, quality of implementation was aimed to be improved. The 13 items of RTOP used in the current study to verify ABI instruction was given in Appendix J.

### 3.7. Data Analysis

In the present study, both quantitative and qualitative data analyses were conducted. Quantitative data analysis was used to address the first three research questions;

while, qualitative data analysis was used to address the fourth and the fifth research questions. The following two sub-sections provide detailed information about quantitative and qualitative data analysis employed in the present study.

### 3.7.1. Data Analysis for the Quantitative Part.

In the quantitative part of the study one group pre-test post-test research design (Fraenkel et al. (2012) was used. More specifically, the effect of the ABI treatment on students' content knowledge, epistemological beliefs and science process skills was examined by comparing students' pre-tests and post-test scores.

In this study, first research question was about uncovering ABI treatment effect on students' content knowledge. There was no control group in this study. Treatment groups' content knowledge was assessed by MHCK and ECK as pre-test and post-test. There were two dependent variables (matter and heat content knowledge and electricity content knowledge) and time was the independent variable representing the implementation of ABI treatment. In line with this, treatment effect on students' content knowledge was analyzed by use of within subject Repeated measure Multivariate Analysis of Variance (MANOVA).

Similarly, second research question was about uncovering ABI treatment effect on students' epistemological beliefs. Epistemological beliefs were measured by use of Epistemological Belief Questionnaire as pre-test and post-test. Epistemological beliefs included three different dimensions which are development, justification and combination of source and certainty. Therefore, there were three dependent variables and one independent variable (i.e., Time) in analysis. Hence, treatment effect on students' epistemological beliefs was analyzed by use of within subject Repeated measure MANOVA.

Likewise, third research question was about uncovering ABI treatment effect on students' science process skills. TIPSII was used to assess students' science process skills. The scale had five dimensions of science process skills. However, reliability scores of these dimensions were found too low in current study. Therefore, analysis was not conducted based on dimensions of science process skills. On the other hand, science process skills were used as one dimensional and students' total scores obtained from pre-test and post-test were used for analysis. In line with this, there was one dependent variable that is science process skills and one independent variable that is time. Therefore, paired sample t-test was conducted to reveal the effect of ABI treatment on students' overall science process skills.

### 3.7.2. Data Analysis for the Qualitative Part.

Qualitative research design was set to answer the last two research questions. Specifically, in qualitative part, the research design involved case-study. Constant comparative analysis was used for analysis. Accordingly, researcher compared a piece of data with another to see similarities and differences in constant comparative analysis. These similarities and differences were grouped and named. Each of these groups was called as category. Then, these categories were used to reveal patterns which address the research questions (Merriam, 2009). How constant comparative analysis was used for the last two research questions is explained in detail below:

Constant comparative analysis was used for research question four aiming to understand $6^{th}$ grade level students' argumentation schemes in the ABI treatment. Student discussions were transcribed, discussion episodes were determined. Accordingly, starting and end points of each discussion episode were determined based on the following criteria: one student proposes his/her idea in a new topic which is the starting point of an episode. Discussion episode lasts until a new discussion emerges in another topic. When another discussion starts, current episode ends. Unit of analysis was reasoning sequence in this study in line with Duschl (2007). Each episode was analyzed deductively in analysis of students' argumentation schemes. Accordingly, Walton's (1996) argumentation schemes and Duschl's (2007) work on argumentation schemes were used for deductive coding to reveal students' argumentation schemes. Table 3.6 presents argumentation schemes, their definitions and sample quotations.

Table 3. 6
*Deductive Codes used to Reveal Students Argumentation Schemes*

| Argumentation Schemes | Description | Example |
|---|---|---|
| Argument from sign | This argument is about students' inferences and their endeavour to explain an observation (Duschl, 2007; Walton, 1996). | The water temperature decreased less in metal cup than plastic cup. We did not expect this, maybe window was open and wind caused more heat loss in plastic cup (Student 111, week 2) |
| Argument from example | This argument is used to support a generalization. The current situation is protected when this argument is used (Walton, 1996). | Solar panels should be established on cities which gets so much sunlight. For example, it can be established in Antalya (Student 144, week 4) |
| Argument from verbal classification | This argument is realized when students' words are unclear. If students clearly tell what they want to say, verbal classification and communication become stronger (Walton, 1996). | When we test the electrical conductivity of matters, bulb brightness was little when we connected cables to the metal cup (Student 223, week 5) |
| Argument from commitment | In this argument; people have their perspectives and they are expected to act based on these perspectives. If there is confliction between their perspective and actions, they are warned and they are asked to behave consistent with their perspective (Duschl, 2007; Walton, 1996). | I want to make counterclaim against student 133. You claimed that you both want to use organic products and chemicals in heat insulation for building. If you are using organic products, this shows that you are environment friendly. Then, it is not meaningful to support use of chemicals if you are environment friendly (Student 111, week 3) |
| Circumstantial argument against the person | The ones using this argument point out conflictions between others speeches and behaviours (Walton, 1996). Due to fact that this argument is highly similar to argument from commitment, this argument was removed from analysis. | This argumentation scheme was not observed. |
| Argument from position to know | In this argument, students ask questions to others when they have limited knowledge. Students look for opposite statement (Duschl, 2007). | When you cover the upside of the cup with aluminium foil, is not it more difficult to melt fat? (Student 223, week 1) |
| Argument from expert opinion | Students use external source or authority to support their claim in this argument (Duschl, 2007; Walton, 1996). | I think we should not use tar in our house for heat insulation because evidence cards say that tar smells bad and we do not want to live in a house that smell bad (Student 231, week 3) |

Table 3.6 cont'd

*Deductive Codes used to Reveal Students Argumentation Schemes*

| Argumentation Schemes | Description | Example |
|---|---|---|
| Argument from evidence to hypothesis | In this argument, there is a testable hypothesis or prediction. Hypothesis is supported or denied based on evidence (Duschl, 2007; Walton, 1996). | Our research question is: which spoon does conduct heat better? We hypothesize that metal spoon is better heat conductor (Student 243, week 1). |
| Argument from correlation to cause | In this inductive argument, student observes a positive relation between two variables, and they think that one is the reason of another although there is no direct observation to support this idea (Duschl, 2007; Walton, 1996). Students prefer plausibility rather than possibility in this argument (Duschl, 2007). | Wood wool can be used for exterior wall as insulating material because wood wool is environmentally friendly (Student 241, week 3). (There was no plausible link between using wood wool for exterior wall and the concept "environmentally friendly") |
| Argument from cause to effect | In this argument, conclusion is directly observed after premise is proposed. There is a clear cause effect relation between premise and conclusion and there is no need for testing (Duschl, 2007) | Aluminium foal was not useful for us because when we made a hole in it using thermometer, there was heat loss from this hole and aluminium could not block heat loss (Student 241, week 1) |
| Argument from consequences | Students consider potential consequences of decisions in this argument. Decisions with good consequences are supported, and decisions with bad consequences are rejected (Duschl, 2007; Walton, 1996). | We should not use glass wool for heat insulation because when we use glass wool, we may suffer to allergic reactions (Student 232, week 3) |
| Argument from comparison | This argument is not specifically mentioned in Walton (1996), but this argument is used when two concrete things are compared according to Walton (1996). | In their experiment, they added four spoons of sugar, but we added one spoon of sugar to test whether sugared water conduct electricity (Student 115, week 5) |
| Argument from analogy | In this argument, students compare two different abstract things (Walton, 1996) | I think that we should make thicker plastic that surround wire. If we make thicker plastic, electricity cannot pass out from thick plastic, there is less heat loss and bulb brightness increases (Student 231, week 6) (In this wrong analogy, student compares flow of electricity with heat conduction) |

Table 3. 6 cont'd

*Deductive Codes used to Reveal Students Argumentation Schemes*

| Argumentation Schemes | Description | Example |
|---|---|---|
| Argument from waste | In this argument, student does not give up her/his argument because s/he made so much effort to construct her/his argument. Students think that if s/he gives up, their effort will be waste, so student keeps his/her argument even though s/he knows it is wrong (Walton, 1996). | This argument was not observed in class. |
| Argument from popularity | In this argument, students support the others and the ideas become major ideas as popular ones (Walton, 1996). | In this experiment, we observed that brightness of bulb increases as response to increasing resistance. As a group, we all think the same thing and we verified our hypothesis (Student 241, week 6) |
| Ethotic argument | When this argument is proposed, students' consider the characters of others to increase validity of their arguments (Walton, 1996). | This argument was not observed in class. |
| Argument from bias | A respondent attack to proponent by thinking proponent has some bias in this argument (Walton, 1996). | This argument was not observed in class. |
| Argument from established rule | In this argument, there is a rule for everyone and everyone should obey the rules (Walton, 1996). In this study, however, argument from established rule was expanded to natural rules that all natural things obey these rules. | I think that we should use solar energy, not wind energy because earth rotates. When earth rotates, one side of the earth is daylight and other side is night. Therefore, we can always benefit from solar energy (Student 221, week 4). |
| Argument from precedent | This argument is refutation of argument from precedent. The rules proposed in established rule argument either modified or refuted (Walton, 1996). In this argument, students proposed limitations of the views supported by established rules. | (In response to argument from established rule: solar energy should be used, but not wind energy, because sunlight always comes one side of the earth)<br>Yes, sun light always come to the world, but you get solar energy in just one side of the earth where you live, therefore you cannot benefit from it at nights and you cannot get it from other side of the world (Student 231, week 4). |

Table 3. 6 cont'd

*Deductive Codes used to Reveal Students Argumentation Schemes*

| Argumentation Schemes | Description | Example |
|---|---|---|
| Argument from gradualism | People are not persuaded by use of one comprehensive argument. Therefore, more than one argument which are linked but less comprehensive are used to persuade others in argument from gradualism (Walton, 1996). | (Explain the relationship between distance of cable ends in salty water and brightness of bulb) Cable that is connected to battery gets current from battery and electricity reach in salty water from the end of this cable. Electricity should pass to other cable's end to light the bulb. When we decrease the distance between ends of cables, more current will pass and less energy will lose in salty water. The more current passes, the brighter bulb we get. |
| The causal slippery slope argument | The causal slippery slope is used to refute argument from gradualism. The weak points of gradualism arguments are used to demolish argument from gradualism (Walton, 1996). | This argument was not observed in class. |
| The precedent slippery slope argument | This type of argument is conservative and does not want to change current situation. The precedent slippery slope argument points out devastating consequences of argument from gradualism. Therefore, if first step of argument from gradualism happens, there will be bad consequences and so first step must not happen (Walton, 1996). | This argument was not observed in class. |
| Argument from vagueness of a verbal classification | This argument is used against argument from verbal classification. This argument claims that there are some vague terms in verbal classification and therefore, argument from verbal classification should not be accepted (Walton, 1996). | You said that there was little brightness of bulb when you connected cable ends to the metal cup. What do you mean by saying little brightness because it is not clear (Student 234, week 5) |
| Argument from arbitrariness of a verbal classification | This argument claims that verbal classification is done arbitrary, and there is no acceptable reason to make this verbal classification so argument from verbal classification should not be accepted (Walton, 1996). | This argument was not observed in class. |

Table 3. 6 cont'd
*Deductive Codes used to Reveal Students Argumentation Schemes*

| Argumentation Schemes | Description | Example |
|---|---|---|
| The verbal slippery slope argument | This argument is used to refute arguments having vague terms. Refutation is done gradually in the verbal slippery slope argument (Walton, 1996). | This argument was not observed in class. |
| The full slippery slope argument | This argument is a rejection to argument from gradualism in order to eliminate negative effects of argument from gradualism. First step of argument from gradualism must not happen. While other slippery slope arguments are related with individuals, the full slippery slope is about public (Walton, 1996). | This argument was not observed in class. |

After each episode in class one was analyzed based on the argumentation schemes given in table 3.6, frequency and percentage of each argumentation schemes were calculated for each week. Bar graphs were used for frequency of argumentation schemes and pie charts were used for percentages.

Fifth research question of this study is "What is the nature of 6$^{th}$ grade students' engagement in argumentation process in ABI treatment?" After student discussions were transcribed, discussion episodes were determined following Sampson and Clark's (2011) approach in their study. After each episode was identified, unit of analysis was determined. Unit of analysis was conversational turns in discussion episode (Sampson & Clark, 2011). Based on unit of analysis and discussion episodes, deductive coding was done to reveal the nature of 6$^{th}$ grade students' engagement in argumentation process. Deductive codes was retrieved from Sampson and Clark (2011) work. Table 3.7 shows deductive codes, their definitions and sample quotations.

All discussion episodes were coded one by one based on these deductive codes for class one in each week. Although data coding included both pre-activity discussions and whole class discussions held after activities, this study just focused on reporting whole class discussion regarding nature of students’ engagement in argumentation process. The reason why pre-activity discussions’ analysis was removed from this study is about absence of evidence. While students used evidence obtained from evidence sheets and their experiments in whole class discussions held after activities, they only used their background knowledge without further evidence in pre-activity discussions. It was thought that lack of evidence found in pre-activity discussion does not accurately reflect students’ engagement in argumentation process so coding of this part was not used in data analysis.

Similarly, small group discussions were not analyzed this study. Actually, it was planned to analyze small group discussions at the beginning of the pilot study. Therefore, tape recorders were brought to class and given some groups in pilot study. However, existence of tape recorders distracted students’ attention and students focused on tape recorder rather than constructing their arguments when tape recorder was used. In line with this, tape recorder was not used in main study and students’ small group discussions were not analyzed. This is one of the limitations of current study.

Table 3. 7

*Deductive Codes used to Reveal Students Engagement in Argumentation Process*

| Code | Definition | Example |
|---|---|---|
| Information seeking | Information seeking is students' questions to others in argumentation process. By using these questions, students request more information or clarification (Sampson & Clark, 2011). | Student 131: In our observations, bulb lighted when we put cables in the salty water. When we decrease the distance among cables, bulb brightness increased. Moreover, water at the end of the cable foamed. A few minutes later water colour turned to yellow...<br>Student 115: *You said that water foamed. How did this happen?* |
| Exposition | Exposition code is about people's own ideas. Proposing ideas in a presentation, clarifying ideas and justifying ideas are examples of exposition (Sampson & Clark, 2011). | Student 144 explains why they covered metal paper clip with aluminium foil (clarifying ideas):<br>Student 144: We think that metal paper clip conducts electricity better if we cover it with aluminium foil because both aluminium and paper clip conduct electricity. If we combine them, the energy passes through materials will be twice. |
| Oppositional Comments | Oppositional comments are students' challenges to others' ideas. When people use oppositional comments they might use evidence, but using evidence is not obligatory to use oppositional comments (Sampson & Clark, 2011). | Student 131 does not accept student 144 explanations given above regarding electric conductivity.<br>Student 131: I think that the energy passes through material will not be twice when you combine paper clip and aluminium foil. Aluminium foil is better conductor, so it decides how much energy will pass. |
| Co-construction of knowledge | Co-construction of knowledge is about the ideas of others. People construct the knowledge together. Co-construction of knowledge includes summarizing, revising, supporting and adding ideas of others (Sampson & Clark, 2011). | Student 223: In our experiment, glass cup lost 25 degree, foam cup lost 19 degree and plastic cup lost 18 degrees, but metal lost 22 degree. Actually, our hypothesis was metal is better heat conductor, but glass lost more temperature in our experiment. Therefore, our claim is glass is better conductor.<br>Student 243: Maybe you touched the glass cup more than others and this caused more heat loss.<br>Student 241: Maybe you kept the thermometer too much and you increased value read on temperature for the metal cup. |

After whole class discussions were analyzed deductively, frequency and percentage of each codes for each week were calculated for class one. These frequencies are presented as bar graphs and percentages are presented as pie graphs in the results chapter. Same deductive coding process was applied to the data obtained from class 2. Then, similarities between class one's coding and class two's coding were recorded consistent with constant comparative analysis. By this way, results common to both classes were identified. Additionally, results which are not consistent across the two groups were determined. The later results were unique to one of the classes. All the findings derived from similarities and differences between two classes can be considered as addressing the research question five which requires the examination of the students' engagement in argumentation process.

Moreover, weekly prepared frequencies and bar graphs for argumentation schemes were compared with weekly prepared frequencies and bar graphs showing students' engagement in argumentation process. By this way, possible connections between nature of students' argumentation process and their argumentation schemes were identified.

## 3.8. Validity

Current study has both quantitative part and qualitative part. According to Fraenkel et al. (2012), validity of a quantitative research includes internal validity and external validity. On the other hand, validity of qualitative research is called as trustworthiness. Trustworthiness of a study is supported by its credibility, dependability, and transferability (Merriam, 2009). Validity of this study considering its quantitative and qualitative parts is presented separately.

### 3.8.1. Validity for Quantitative Part.

Quantitative part of this study involved one group pre-test post-test design where there is no control group. It is one of the weak experimental designs (Fraenkel et al, 2012) Internal and external validity of the quantitative part, therefore, are explained considering the properties of this design.

#### *3.8.1.1. Internal Validity.*

According to Fraenkel et al. (2012), internal validity means that the observed differences on dependent variable is the direct result of independent variable and other factors should not affect dependent variable. Quantitative part of this study involved one group pre-test post-test design which unfortunately has many internal validity threats because of the absence of control group (Fraenkel et al., 2012). According to Fraenkel et al. (2012), one group pre-test post-test experimental design suffers nine internal validity threats. In other word, nine different reasons, other than the treatment, might cause change in the dependent variables. These possible causes are history, maturation, instrument decay, data collector characteristics, data collector bias, testing, statistical regression, attitude of subjects and implementation. Accordingly, how each factors threat the internal validity and how researcher took actions to avoid these threats are presented below.

History effect is unplanned or unanticipated event that happens during the course of study (Fraenkel et al., 2012). In this study, no such events occurred. In addition, to prevent disruption of the implementation do to unexpected events, the schedule of the study was well planned prior to the study and shared with the teachers and school administration.

In this study, maturation is not considered as a potential threat either, because maturation becomes a threat if the change on dependent variable is caused by passing of time rather than the treatment (Fraenkel et al., 2012). Duration of the current study was six weeks and six weeks is a relatively short period of time. Therefore, it is thought that passing of time in this short term most probably do not affect the dependent variable.

Another potential threat to internal validity in the current study was instrument decay which is about changing nature of instrument. If instrument is open to different interpretation or it is too long for assessing, instrument decay threat may emerge (Fraenkel et al., 2012). Instrument decay is not considered as threat in this study because the instruments used were either in multiple choice item formats or consisted of Likert type items which are completely objective to score, so they were not open to different interpretations. Thus, the use of objective items with easy scoring eliminated possible threat to internal validity.

Data collector characteristics such as gender, age, ethnicity, and language pattern can also affect the nature of data (Fraenkel et al., 2012). In addition, data collector may unconsciously distort the data to support hypothesis (data collector bias) (Fraenkel et al., 2012).The quantitative data obtained in this study was collected by objective instruments such as multiple choice test items and Likert type scales and the data was coded in SPSS program by researcher exactly as the same way that students marked. Therefore, it is thought that characteristics of researcher (data collector) and data collector bias do not pose an internal validity threat in the present study.

The use of pre-test can affect students' performance on post-test and this threat is known as testing threat For example, students are alerted to practice because of the items found in pre-test, and they seek the answers of pre-test searching textbook and other sources. This increased effort may increase their performance on post-tests (Fraenkel et al., 2012). Although this threat was not totally eliminated in this study, researcher told students that these treatment, pre-tests, or post-tests did not affect their grades. Students were asked to answer each question honestly. By this way, researcher wanted to normalize the conditions. This normalization might decrease the testing effect, but still there is a possibility that pre-tests affected students' performance because the same questions were asked in post-tests and students could remember these questions from pre-tests.

Regression is another threat and existence of people having extreme characteristics related with tests can cause regression threat. Accordingly, if a person's pre-test score is too low, most probably it will improve at the end of the study regardless of the effect of independent variable. In the same manner, some scores may be too high in pre-test, and these scores cannot improve in post-tests even though implementation is successful (Fraenkel et al., 2012). Prior to the study, classes having average achievement were included to the study. By this way, possible extreme scores were aimed to be eliminated.

Attitude of subjects, which is related to how subjects perceive the study and their involvement in it, is another potential threat to internal validity in this study. Participants may think that treatment is something special and researcher cared them. Therefore, participants might improve their performance to assist researcher (Fraenkel et al., 2012). At the beginning of the treatment, researcher told subjects that this treatment is part of their education. In this way, it was aimed to eliminate students' perceptions that they are getting a type of distinctive intervention.

Regarding implementation as a potential threat, Frankel et al. (2012) suggest that researcher should not be implementer to avoid implementation threat because researchers might unintentionally increases the effect of treatment. In line with this, teacher implemented the lessons and researcher used a checklist to control whether teacher implement the treatment accurately. By this way, treatment effect was not exaggerated or inhibited.

After describing the possible threats to internal validity and researcher's precautions to avoid these threats, next part presents information about external validity of quantitative part.

#### *3.8.1.2. External Validity.*

External validity concerns to what extent findings of the study can be applied to other situations (Merriam, 2009). While results of a treatment is generalized, researchers should focus on three possible external validity threats which are interaction of selection and treatment, interaction of setting and treatment, and interaction of history and treatment (Creswell, 2009).

Interaction of the selection and treatment threat refers that participants' characteristics are narrow; therefore, researcher could not generalize the findings to other individuals who do not have these characteristics. Creswell (2009) suggests that researchers should point out group characteristics and claim that generalizations of these groups' scores cannot be done other groups not having common characteristics. For example, current study was held with 6th grade level, 11-12 age years' old students having average achievement level. Therefore, findings of the study can be generalized to other groups having same characteristics. However, findings may not be generalized to other students from different grade level, age and achievement level. Future studies can be held with students having other characteristics that subjects of this study did not have. By this way, external validity of this research can be supported and interaction of selection and treatment threat can be avoided.

Interaction of setting and treatment threat shows that setting of a study might inhibit generalization of the findings to other settings (Creswell, 2009). Context of the study was described in sample and population part. Accordingly, this study was conducted in one of the schools located in a central district of Ankara. The school was a public school and students' families' socio-economic status were high. Therefore, findings of this study can be generalized to other similar contexts. On the other hand, findings may not be generalized to other context such as private schools, distant towns from the city and places where socioeconomic status is not high. For example, Akkus et al., (2007) reported that students from low achievement group benefitted from ABI instruction more than students from high achievement group. It is possible that proportion of low achievement students would change (most probably increase), if this study was conducted in distant towns or the places where students' families have low socioeconomic status. The treatment effect might increase in these disadvantageous places. Therefore, findings of current study may not be generalized to other contexts. Future studies can be done in other contexts to increase external validity of the study and to eliminate interaction of setting and treatment threat.

Interaction of history and treatment threat is about the time when treatment is conducted. Accordingly, findings of the treatments cannot be generalized to past and future time (Creswell, 2009). This study was held in spring semester of 2018-2019; therefore, findings are generalized in this time period. For example, Dindar and Taneri (2011) reported that Turkish science curriculum program emphasized content knowledge and observation in 1968. On the other hand, this program ignored role of laboratory and students’ active participation to science lessons in those years. Therefore, findings of the current study cannot be generalized to those years. On the other hand, nowadays MONE (2018) suggests that students should actively engage in scientific argumentation, conduct their own experiments, and construct knowledge. Therefore, findings of the study can be generalized to other schools following suggestions of MONE (2018). The repetition studies can be done to generalize findings for the future to avoid this threat.

Up to this point, validity of quantitative part of this research was explained. The chapter continues with validity of qualitative part of research.

### 3.8.2. Validity for Qualitative Part (Trustworthiness).

Qualitative part of the study includes observations of two classes which were done to reveal students’ engagement in argumentation process and their argumentation schemes. Hence, validity of qualitative part focuses on these observations. Nomenclature of validity in qualitative research is different from quantitative research. Validity is named as trustworthiness in qualitative research and trustworthiness includes credibility, dependability, transferability (Lincoln & Guba, 1985). The following parts inform trustworthiness of the study in detail.

#### *3.8.2.1. Credibility.*

Credibility corresponds to internal validity issue in quantitative research and it refers to the consistency between research findings and reality. Merriam (2009) claims that credibility of a qualitative research increases using triangulation, member check, adequate engagement in data collection, researcher position (reflexivity), and peer review.

According to Merriam (2009), use of multiple methods of data collection, data sources, investigators, and theories support triangulation. Multiple methods of data collection were not used in this study; the only method used by researcher was observation. Other methods like interview could have accompanied with observations, but it did not. This can be considered as one of the limitations of current research. On the other hand, multiple data sources were used in current study. Classroom observations were done in two classes throughout the study. Observations held in different weeks and different classes were transcribed, compared and contrasted. These multiple data sources contributed the triangulation. Another possible data source used in this study was student templates. However, student templates inform about students' written arguments rather than their oral argumentation. On the other hand, focus of this study is students' oral argumentation. Therefore, student templates were not used as data source in this study. Likewise, data found in student template was not analyzed. On the other hand, student template was used as scaffolding that assist students to construct arguments and engage in argumentation process in this study. Likewise, multiple investigators involved in the study increasing triangulation. Accordingly, research team members shared their ideas about implementation throughout the research. Their feedbacks were used to improve the research quality. Likewise, one of the research members contributed to the selection of data analysis method. There have been several conversations between research team members increasing researcher and research mates' cumulative knowledge on argumentation. Furthermore, one of the research members participated in qualitative data analysis. All in all, multiple investigators made many contributions to this study. Finally, it can be claimed that multiple theories were used to increase triangulation. These multiple theories are use of both argumentation process and argumentation schemes in this study. It is thought that argumentation process and argumentation schemes complete each other. Accordingly, if this study had only included argumentation process, there would not have been any suitable inference regarding the product of argumentation. Likewise, if this study had only used argumentation schemes, there would not have been any information regarding argumentation process. Therefore, use of both argumentation process and argumentation schemes in qualitative part as multiple theories contributed triangulation of the study. On the other hand, member check was not done in this study because students were video-recorded and their speech and what they implied were clear, so member check was not needed. Moreover, data was analyzed with one of the research mates and students' speech and these speeches' meanings were discussed with this research mate. By this way, what students told in their argumentation was further clarified.

Another way of increasing credibility of a research is adequate engagement in data collection (Merriam, 2009). Researcher spent so much time with participating teacher and visited the school more than two years during pilot and main studies. Specifically, there were several conversations held with participating teacher in both pilot and main studies. Likewise, researcher observed the full implementation of the ABI treatment in both pilot and main studies. Moreover, researcher trained both teacher and students prior to research. All these activities that researcher participated in provided adequate engagement in data collection. Through these conversations, trainings and observation of class instances; researcher got familiarity with school contexts, students, and the teacher.

Another way of increasing credibility is researcher position or reflexivity. According to Lincoln and Guba (2000), human being is the instrument in qualitative study. Therefore, researcher should critically examine herself or himself to what extent s/he is sufficient to conduct current research. One factor affecting reflexivity is researcher's experience. First of all, researcher has experience in qualitative research: the researcher conducted qualitative research in master thesis. Likewise, researcher has qualitative research experiences because researcher collected data in university's scientific research projects. Moreover, researcher had articles on science education journals and all these articles included qualitative research design. It is thought that researcher is not only capable of conducting qualitative research, but also has experience on argumentation research. Accordingly, researcher enrolled in an argumentation course in 2012. In this course, researcher got some theoretical and practical experience about argumentation. In 2016, researcher prepared a proposal for this study and researcher reviewed the literature which contributed to researcher's theoretical understandings. In 2017, researcher prepared 29 lesson plans using argumentation as teaching method. Nine different argumentation approaches were used in these lesson plans. These approaches are attention grabbing approach (Wojdak, 2010), analogical mapping based comparison activities (Emig, McDonald, Zembal-Saul, & Strauss, 2014), use of multiple representations (Namdar & Shen, 2016), coupled inquiry cycle (Dunkhase, 2003), history and philosophy of science (Archila, 2015), hypothetico-predictive argumentation (Lawson, 2003), simulated jury (Vieira, Bernardo, Evagorou, & Melo, 2015), argument-based inquiry (Hand et al., 2004), and modified argument driven inquiry (Chen, Wang, Lu, Lin & Hong, 2016). While writing these lesson plans, researcher got familiarity to argumentation. Likewise, researcher read and summarized many argumentation articles and books to better understand its philosophy and practical stance in science education. Furthermore, researcher planned and conducted pilot study of this research in 2018. This process added contribution to researcher's understanding about its theoretical underpinnings and challenges of using argumentation in practice. Finally, researcher planned, prepared, conducted the main study in 2019. All of these activities that researcher engaged in added him experience about argumentation supporting researcher's position to conduct this activity.

Credibility of the study was also supported with peer review. Accordingly, researchers having experience in methodology and content add contribution to study as peer review (Merriam, 2009). Two experts, one associate professor and one assistant professor in science education field, provided feedback to researcher throughout the study. Moreover, two professors in science education monitored the whole process. They had conversations with researcher and shared their suggestions. Likewise, researcher met with research team members and discussed with them about their suggestions to improve quality of research. Furthermore, one members of research team participated in qualitative data analysis process. It is thought that all these reviews, feedbacks, conversations and suggestions as peer reviews increase credibility of current research. Next part continues with dependability of the study.

#### *3.8.2.2. Dependability.*

Dependability or consistency is about reliability of the study. Results and collected data should be consistent in qualitative research (Merriam, 2009). Data was coded by researcher and one of the research team mates regarding students' engagement in argumentation process and argumentation schemes. Prior to data analysis, researchers coded first week's data together in order to make calibration. By this calibration, researchers established a common understanding such as negotiating on standards for coding. After calibration, analysis was done separately. After researchers analyzed the data they compared the results of data analysis. Then, inter-rater agreement was calculated based on the formula given by Miles and Huberman (1994):

Inter-rater agreement = (Number of Agreement) / (Number of agreement + disagreement)

Inter-rater agreement for argumentation process was calculated as 85 % and inter-rater agreement for argumentation schemes were calculated as 80 %. After researchers detected different codes, they talked about them and discussed until conflictions were solved. Following is an example of conflictions between coders regarding students' engagement in argumentation process:

Student143: When we increase number of bulbs, the bulbs have to share energy provided by battery, so bulb brightness decreases.

Teacher: Do these bulbs have to share energy?

Student 143: Yes.

Teacher: Why?

Student 143: because both bulbs are in same circuit and battery energy passes through the conductive wires and each of the bulbs.

In this example, researcher coded student143 speech as co-construction of knowledge. On the other hand, other coder thought that this is just the ideas and answers of student143 and there is no co-construction. Coder added that there should be some other ideas supporting student143 to code this speech as co-construction. Therefore, code of this dialog should be exposition. This dialog persuaded researcher and this excerpt was re-coded as exposition, not co-construction. Other conflictions between coders were resolved in same way. Next part provides information about transferability of the study.

#### *3.8.2.3. Transferability.*

Transferability is about the external validity of qualitative study. It considers to what extent findings can be applied to other studies (Merriam, 2009). Although findings of qualitative part cannot be generalized to other studies, researchers having similar contexts including 6$^{th}$ grade students enrolled in public schools located in central districts of metropoles can benefit from the findings of the study. According to Merriam (2009), transferability of research can be supported by rich thick description of the study and maximum variation. Rich thick description of the study was done through explanation of research design, sample and participants part, teacher training, pilot study, main study (i.e., treatment) and excerpts obtained from observations. Likewise, findings of qualitative part were prepared based on observation of two classes. Use of two classes as data sources provided maximum variation. If this study had included only one class, inferences obtained from observations would have been limited with what this class presented. On the other hand, use of two classes enriched the data and provided more inferences by maximizing the variation. In this part validity of qualitative part (trustworthiness) of the study was explained based on credibility, dependability and transferability. Next part explains ethical considerations of the study.

## 3.9. Ethics

After instruments and treatment were prepared, application was done for university's ethical committee approvals for the pilot study in 2018. After obtaining university approval, application was done for Ministry of National Education (MONE) ethical committee approvals. After getting MONE's approval, permissions were taken from school principal. Then, teacher and students voluntarily participated in pilot study. In this process, participation form was distributed the students. This form provided brief information about study, participants' rights, and researcher's information. Furthermore, parent permission form was prepared and sent to families to get their permissions about children's participation to pilot study. Families accepted participation of their children in to study. Teacher communicated with parents in this process. Content of the parent permission form included brief information about researcher, information about study, students' right, and parents' rights. Same process regarding ethical permissions was repeated for main study in 2019. University's ethical committee approval is presented in Appendix K. Likewise, Ministry of National Education's ethical committee approval is presented in Appendix L. Participation form that was distributed to students are given in Appendix M and parent permission form can be found in Appendix N.

Both teacher's and students' right were protected in this study. Teacher, parents and students had rights that they can leave the study whenever they want. The real names were not used in this study. Labelling of each student was explained in sample and participants part. Likewise, data were not shared with third person. Researcher shared data only with experts to get their feedbacks and analyze data. None of the participants were deceived or no one got harmed in this study. Furthermore, researcher did not judge teacher or students throughout the study and so participants felt themselves in a safety environment. In a similar vein, students got familiarity with researcher and existence of camera; therefore, existence of researcher and camera did not affect students' participations and did not result in ethical problems.

Data analysis can also include ethical problems. For example, researchers might select some of the data reflecting their expectations and ignore anomalous data. In this study, all data obtained from whole class discussions were analyzed in same way and none of the data was left without analysis. Likewise, data was not distorted in analysis process. Researchers adhered to codes and their meaning during analysis. Next part informs regarding the limitations of the study.

### 3.10. Limitations

Current study has some limitations. Regarding quantitative part, convenient sampling was used in this study. In other word, a group of students participated in current study because they were available. If random sampling had been used to collect data, it would have been better to generalize the findings of the study (Fraenkel et al., 2012).

Current study has also three limitations regarding its qualitative aspect. First limitation is about data analysis. Qualitative data analysis of this study included whole class discussions done after experiments and activities. However, data obtained from pre-activity discussion held prior to main activity were not analyzed. The reason why these discussions were not analyzed was absence of evidence. Students only used their background knowledge in these discussions and they did not use experiments and evidence sheets which are more scientific than their background knowledge. If these pre-discussions had been included in data analysis, nature of their argumentation process might have been portrayed wrong because arguments would have been constructed on naïve ideas rather than scientific evidence. Similarly, small group discussions were not analyzed in this study as another limitation regarding data analysis. In pilot study, students over reacted to existence of sound recording and they did not work properly when sound records were provided them. Therefore, main study did not include sound record and absence of data for small group discussions caused not to analyze small group discussions. If small group analysis had been added to current study, it would have been better represent students' engagement in argumentation process. However, students' engagement in argumentation process is limited with whole class discussions in current study. Another limitation of the study is about use of multiple methods of data collection for qualitative research. Observation was the only methods of data collection used to gather students' engagement in argumentation process. If it has been combined with interviews, more information about argumentation practice would have been gathered. Final limitation of the study is about transferability of findings. Researchers having similar participants and context can benefit from the findings of the study. For example, findings of current research are efficient for research conducted with 6$^{th}$ grade students enrolled in public schools located in metropolis. However, findings may not be beneficial for other ABI research conducted with different grade levels, different type of schools like private schools and remote areas from the cities. After limitations of the study, assumptions of this research are presented in the following part.

### 3.11. Assumptions

There are three main assumptions accompanying this study: first, it is assumed that the students honestly and seriously responded to the instruments used in the quantitative part. Secondly, it is thought that students' engagement in argumentation process was not affected by existence of researcher and use of video cameras in qualitative part. Thirdly, research team members assisted students when students had difficulty to conduct their investigation that means team members facilitated and stimulated students' work. Therefore, it is assumed that assistance of team members did not affect the results.

# CHAPTER 4

# RESULTS

This chapter reports results for research questions, and so this chapter includes five parts. First part presents results about changes of participants' content knowledge (heat and matter and electricity). Second part presents results about changes of participants' epistemological beliefs (e.g. source and certainty of knowledge). Third part shows the results about changes of participants' science process skills (i.e. integrated process skills). Fourth part informs about argumentation schemes that 6$^{th}$ grade students used throughout the study. Last past demonstrates students' engagement in argumentation process.

## 4.1. Results for Changes in Content Knowledge

First research question of the study was "Is there a change in 6$^{th}$ grade students' content knowledge from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?" Content knowledge included two units namely matter and heat unit and electricity unit. Therefore, there were two dependent variables which are matter and heat content knowledge and electricity content knowledge and one independent variable which is time (Time 1 and Time 2). In line with this, within subject repeated measures multivariate analysis of variance (MANOVA) was planned to be conducted. Before inferential statistics, results about descriptive statistics are presented.

### 4.1.1. Descriptive Statistics for Changes in Content Knowledge.

Table 4.1 presents participants' descriptive statistics of content knowledge for pre-test and post-test. Participants' matter and heat content knowledge was assessed by MHCK test. Descriptive statistics suggested that participants' matter and heat content knowledge mean score increased from 9.77 (SD=3.42) to 11.65 (SD=4.15) from pre-test to post-test. Minimum and maximum scores for matter and heat content knowledge were 3 and 16, so range was 13 in pre-test. On the other hand, minimum and maximum values increased to 4 and 18 in post-test (Range=14). Likewise, participants' electricity content knowledge was assessed by ECK test. Accordingly Table 4.1 participants' electricity score improved from 9.39 (SD=4.37) to 11.94 (SD=4.26) from pre-test to post-test. Minimum and maximum values for electricity content knowledge was 1 and 19 (Range=18) in pre-tests and these values were 3 and 19 (Range=16) in post-tests. The next part informs results about inferential statistics.

Table 4. 1

*Descriptive Statistics for Content Knowledge*

| | Pre-Test | | | | | | Post-Test | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variable | N | Mean | SD | Range | Min. | Max. | N | Mean | SD | Range | Min. | Max. |
| MHCK | 71 | 9.77 | 3.42 | 13 | 3 | 16 | 71 | 11.65 | 4.15 | 14 | 4 | 18 |
| ECK | 70 | 9.39 | 4.37 | 18 | 1 | 19 | 70 | 11.94 | 4.26 | 16 | 3 | 19 |

Prior to Repeated measure MANOVA; its assumptions are tested to understand to what extent the results yield to valid interpretations. Therefore, firstly assumptions of the repeated measure MANOVA are tested and secondly, results of the repeated measure MANOVA is presented to answer research question 1 in this part.

### 4.1.2. Evaluation of the Assumptions of Repeated Measures Multivariate Analysis of Variance for Content Knowledge.

Seven assumptions were considered before doing within subject multivariate analysis of variance test. These assumptions are missing data analysis, univariate normality, univariate outliers, multivariate normality, multivariate outliers, linearity and absence of multicollinearity and singularity. These assumptions are discussed separately in this part.

#### *4.1.2.1. Missing Data Analysis.*

This study included 71 students however; two of the students either did not complete pre-tests or post-test. Thus, the missing data were less than 3 %. According to Tabachnick and Fidell (2012), if the percentage of missing data is low (i.e. 5 % or less) in a random pattern, nearly all missing data handling procedures yield comparable results. Considering this idea and the nature of the design of this study (i.e. pre-test post-test design), the data from these students were excluded and the remaining data from the 69 students were used to conduct the analysis.

According to Hahs-Vaughn (2017), product of the number of dependent variables and number of measurement occasions that is the number of within subject factors should be less than sample size. In this part of the study, there are 2 dependent variables (e.g. electricity content knowledge) and each was measured 2 times (e.g. pre-test). Sample size is larger than the product of these values. When the number of dependent variables and number of measurement occasions was taken into consideration, this remaining sample size was large enough to conduct the analysis

#### *4.1.2.2. Univariate Normality.*

Statistical or graphical methods are used for the assessment of normality (Tabachnick & Fidell, 2012). Skewness and kurtosis values are considered to assess normality. While skewness represents the symmetry of distribution, kurtosis is about whether a distribution is peaked or flat. Normal distributions have zero value for skewness and kurtosis (Tabachnick & Fidell, 2012). Therefore, skewness and kurtosis values were used to assess univariate normality assumption. Table 4.2 shows skewness and kurtosis values regarding matter and heat and electricity content knowledge;

Table 4. 2

*Skewness and Kurtosis Values for Content Knowledge*

| | Skewness | Kurtosis |
|---|---|---|
| Pre-test | | |
| MHCK | 0.09 | -0.67 |
| ECK | 0.10 | -0.85 |
| | | |
| Post-test | | |
| MHCK | -0.24 | -1.09 |
| ECK | -0.04 | -1.04 |

According to Hahs-Vaughn (2017), distributions having skewness values larger than 2 and kurtosis values beyond 7 are considered as non-normal. Table 4.2 suggested that skewness values for test scores are between -0.24 and 0.10. Likewise, kurtosis values are between -1.09 and -0.67. There is no skewness and kurtosis value that extend critical values. Therefore, it can be claimed that univariate normality assumption is met.

#### *4.1.2.3. Univariate Outliers.*

“Univariate outliers are cases with an outlandish value on one variable” (Tabachnick & Fidell, 2012, p.73). If a score has too large standardized value (z-score), this score is univariate outlier. Standardized scores which are beyond +3.29 and -3.29 are potential univariate outliers. Therefore, students’ scores were standardized in order to detect potential univariate outliers. Table 4.3 presents highest and lowest scores for each test:

Table 4. 3

*Standardized Scores for Content Knowledge*

| | Highest Z-scores | Lowest Z-scores |
|---|---|---|
| Pre-test | | |
| MHCK | 1.88 | -1.96 |
| ECK | 2.23 | -2.00 |
| | | |
| Post-test | | |
| MHCK | 1.56 | -1.82 |
| ECK | 1.64 | -2.10 |

As it is seen from Table 4.3, highest and lowest z-scores for matter and heat and electricity pre-tests and post-tests ranged between +2.23 and -2.10. These values are between critical values which are +3.29 and -3.29. Therefore, it can be claimed that there is no univariate outliers in cases.

#### *4.1.2.4. Multivariate Normality for Dependent Variables.*

According to Hahs-Vaughn (2017), linear combination of variables should be normally distributed for multivariate normality for dependent variables. Prior to multivariate normality, first univariate normality assumption should be met. Univariate normality was met and explained in previous parts. However, univariate normality is not sufficient for multivariate normality (Hahs-Vaughn, 2017).

There is no direct way to measure multivariate normality assumption in MANOVA. Therefore, this assumption was not directly tested. Moreover, Hahs-Vaughn (2017) claimed that MANOVA including repeated measure MANOVA is robust to violations of multivariate normality when there are 20 or more cases per cell. In this study, 69 students completed both tests in pre-test and post-test. Therefore, violating this assumption is not serious.

Moreover, multivariate outliers can negatively affect multivariate normality. In other word, existence of multivariate outliers might violate multivariate normality assumption (Hahs-Vaughn, 2017). The next part informs whether multivariate outliers exist in analysis.

#### *4.1.2.5. Multivariate Outliers.*

"Multivariate outliers are cases with an unusual combination of scores on two or more variables" (Tabachnick & Fidell, 2012, p.73). Mahalanobis distance is used to detect multivariate outliers and Mahalanobis distance is "the distance of a case from the centroid of the remaining cases where the centroid is the point created at the intersection of the means of all the variables." (Tabachnick & Fidell, 2012, p.74). Accordingly, Mahalanobis distance of cases is compared with the value obtained from chi square $\chi^2$ table. In this case, number of dependent variable is two and so degree of freedom is also two and corresponding chi-square value is 13.82 (Pallant, 2007). Similarly, cases having p value less than 0.001 for this chi square value are accepted as outlier (Tabachnick & Fidell, 2012). Mahalanobis distance for each case was calculated by conducting linear regression analysis in SPSS 26 program. After Mahalanobis distances were calculated, cases were sorted in descending order and cases having the highest 10 Mahalanobis distance are reported in Table 4.4. Moreover, using this Mahalanobis distances, a new variable was defined to calculate corresponding p-values for each Mahalanobis distance. 1- CDF.CHISQ (MAH_1, 2) code was used to calculate p-values. These p-values are also available in Table 4.4:

Table 4. 4

*The Highest 10 Scores for Mahalanobis Distance for Content Knowledge*

| Highest Scores | Mahalanobis Distance | Probability level |
|---|---|---|
| 1 | 13.11 | 0.0014 |
| 2 | 5.63 | 0.0598 |
| 3 | 4.43 | 0.1092 |
| 4 | 4.18 | 0.1234 |
| 5 | 4.18 | 0.1238 |
| 6 | 3.53 | 0.1709 |
| 7 | 3.42 | 0.1803 |
| 8 | 3.38 | 0.1840 |
| 9 | 3.37 | 0.1856 |
| 10 | 3.34 | 0.1879 |

The highest Mahalanobis distance is 13.11. This value does not extend the critical values (13.11<13.82). Likewise, smallest p value is 0.0014 and this value is larger than critical value (p>0.001). Therefore, it can be claimed that there is no multivariate outlier in data.

#### *4.1.2.6. Linearity.*

Linearity assumption is provided when there is a linear correlation between variables (Tabachnick & Fidell, 2012). Bivariate scatterplots between pairs of variables was used to detect possible non-linearity. Matrix scatterplot to check linearity assumption is presented in Appendix O. According to Matrix scatterplot, there are linear relationships between variables and there is no sign for non-linearity. Therefore, it can be concluded that linearity assumption was met.

#### *4.1.2.7. Absence of Multicollinearity and Singularity for content knowledge changes.*

According to Tabachnick and Fidell (2012), multicollinearity problem occurs when the correlation between variables are very high ($r>0.90$). Multicollinearity problem shows that variables are too similar. Correlation coefficient between pairs of variables was used to detect any multicollinearity problem. Table 4.5 shows correlation coefficient among variables:

Table 4. 5

*Corelation Coefficient Values for Content Knowledge*

| | Matter and Heat Pre-test | Matter and Heat Post-test | Electricity Pre-test | Electricity Post-test |
|---|---|---|---|---|
| Matter and Heat Pre-test | | 0.58 | 0.68 | 0.54 |
| Matter and Heat Post-test | 0.58 | | 0.68 | 0.72 |
| Electricity Pre-test | 0.68 | 0.68 | | 0.61 |
| Electricity Post-test | 0.54 | 0.72 | 0.61 | |

As seen from Table 4.5, correlation coefficient values are between 0.54 and 0.72. These values are not higher than critical value (.90), therefore, it can be said that correlation between variables are not too high and variables are not similar, hence it can be concluded that there is no multicollinearity problem in data.

Second concern is singularity and this problem happens when one variable is combination of others (Tabachnick & Fidell, 2012). In this test, there are two dependent variables which are matter and heat content knowledge and electricity content knowledge. Tests used to assess these variables were constructed separately and their core ideas and objectives are different, thus, it can be claimed that one variable do not comprehend another and there is no singularity problem.

In conclusion, data obtained from matter and heat content knowledge and electricity content knowledge tests has no serious violations for the repeated measure MANOVA assumptions. Therefore, repeated measure MANOVA test can be conducted to reveal the ABI treatment effect on students' content knowledge.

### 4.1.3. Repeated Measures of Multivariate Analysis of Variance for Content Knowledge.

Repeated measures MANOVA was conducted to reveal content knowledge changes as response to ABI treatment. Time was the independent variable and there were two dependent variables which are matter and heat content knowledge and electricity content knowledge in this test. Main test results are presented in Table 4.6. Results showed that there was a statistically significant effect of argument-based inquiry treatment on combined dependent variables, $F(2,67)= 20.55$, $p= 0.00$; Wilks' Lambda= 0.62.

Table 4. 6

*Multivariate Test Result within Subjects for Content Knowledge*

| Effect | Wilks' Lambda | F | Hypothesis df | Error df | Significance Level |
|---|---|---|---|---|---|
| Time | 0.62 | 20.55 | 2 | 67 | 0.00 |

Paired sample-t tests were conducted as follow up tests to understand which content knowledge (e.g., electricity) statistically changed as response to treatment. However, it should be noted that paired sample-t test does not consider the correlation between dependent variables and so it is not powerful.

Due to fact that there are two tests (pair sample t-test for matter and heat, pair sample t-test for electricity), α level was set as 0.25 (0.50/2) and confidence interval was determined as 97.5% prior to analysis. Paired sample t-test analysis results are presented in Table 4.7.

Table 4. 7

*Paired Sample t-test Results for Content Knowledge*

| Pairs | t | Df | Sig. | Mean Difference | Confidence Interval (97.5 %) Lower | Confidence Interval (97.5 %) Upper | Eta Squared |
|---|---|---|---|---|---|---|---|
| Matter and Heat post-Matter and Heat pre | 4.48 | 68 | 0.000 | 1.90 | 0.93 | 2.87 | 0.23 |
| Electricity post-Electricity pre | 5.47 | 68 | 0.000 | 2.46 | 1.43 | 3.50 | 0.31 |

According to the paired sample t-test results given in Table 4.7, there was a statistically significant increase in matter and heat content knowledge scores from pre-test ($M=9.64$, $SD=3.37$) to post-test ($M=11.54$, $SD=4.13$), $t(68)=4.48$, $p<0.025$ (two tailed). The mean difference in matter and heat content knowledge scores was 1.90 with a 97.5 % confidence interval ranging from 0.93 to 2.87. The eta squared statistic (0.23) indicated large effect size according to Cohen (1988).

The following figure provides more information regarding how students' matter and heat content knowledge changed from pre-test to post-test.

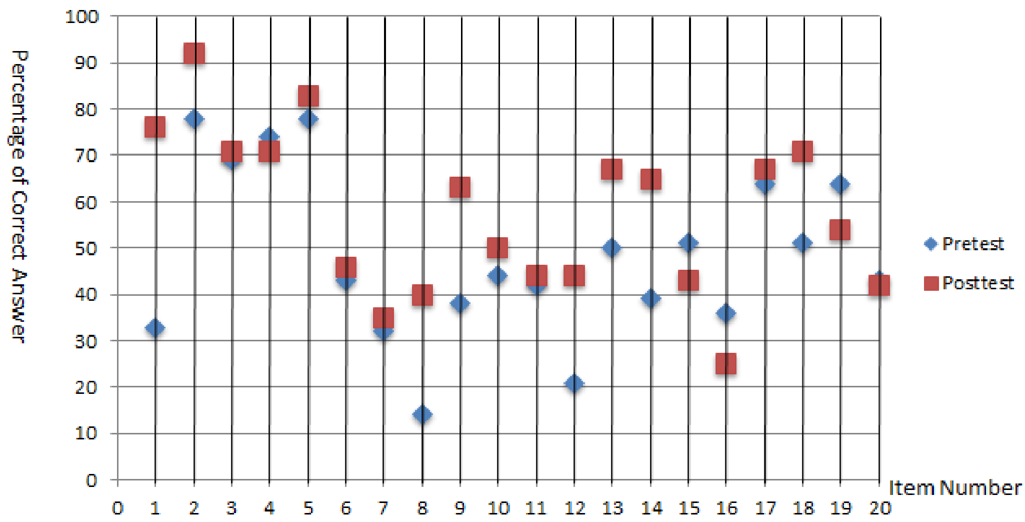



*Figure 4. 1* Comparison of pre-test and post-test scores regarding matter and heat knowledge

Red square represents percentage of correct answer for given item in post-tests, and blue square (diamond shape) represents percentage of correct answer for given item in pre-tests. According to Figure 4.1, percentages of post-test scores are higher than percentages of pre-test scores regarding MHCK. Students' content knowledge increased in 15 items out of 20 from pre-test to post-test. On the other hand, students' content knowledge decreased in 4 items namely item 4, item 15, item 16 and item 19. Similarly, students' performance did not change in last item (item 20). Actually, there was no so much difference in item 4, and this question was about heat transfer topic taught at the beginning of the treatment (week 1). It is possible that students did not adapt to study at the beginning, hence their knowledge about this item might not increase. Similarly, students' performance did not improve in item 15. In this item, students were expected to compare disadvantages of renewable and non-renewable energy sources in terms of economy and environment. Moreover, students were expected to interpret data found in item 15 to compare different energy sources. It is possible that both interpreting table, comparing disadvantages of different sources, and connecting this interpretation and comparison are difficult task for students and because of this reasons, students might not improve themselves in item 15. Next, item 16 was about biomass as source of renewable energy. While solving this item, students were expected to think from two opposing perspective to reach correct answer. First, they should know that biomass contaminates the air after it is burnt. Second, they should grasp the idea that biomass makes photosynthesis and this process leads them to clean air. Students might have difficulty to think these two opposing idea and linking them, therefore, they might not successfully answer this item in post-test. Likewise, students could not improve their performance in item 19. This item was about interpreting knowledge given in table, fossil fuels' damage to environment, fossil fuels' damage to human beings. Moreover, this question asked students to link fossil fuel damages to environment and human beings. These four requirements might be compelling for students. All these four items were in analysis and synthesis level. Likewise, students' performance did not change in item 20 and this question was in evaluation level. On the other hand, students improved their matter and heat content knowledge other 15 items. Six of these 6 items were in knowledge level, 5 of them were in analysis level, 2 of them were in synthesis level, one of them was in application level and other one of them was in evaluation level. In conclusion, it can be said that participants' matter and heat content knowledge increased in this study.

Likewise, there was a statistically significant increase in electricity content knowledge scores from pre-test ($M=9.52$, $SD=4.25$) to post-test ($M=11.99$, $SD=4.28$), $t\ (68) = 5.47$, $p<0.025$ (two tailed). The mean difference in electricity content knowledge scores was 2.46 with a 97.5 % confidence interval ranging from 1.43 to 3.50. The eta squared statistic (0.31) indicated large effect size according to Cohen (1988). Figure 4.2 provides detailed information about participants' change on electricity content knowledge.

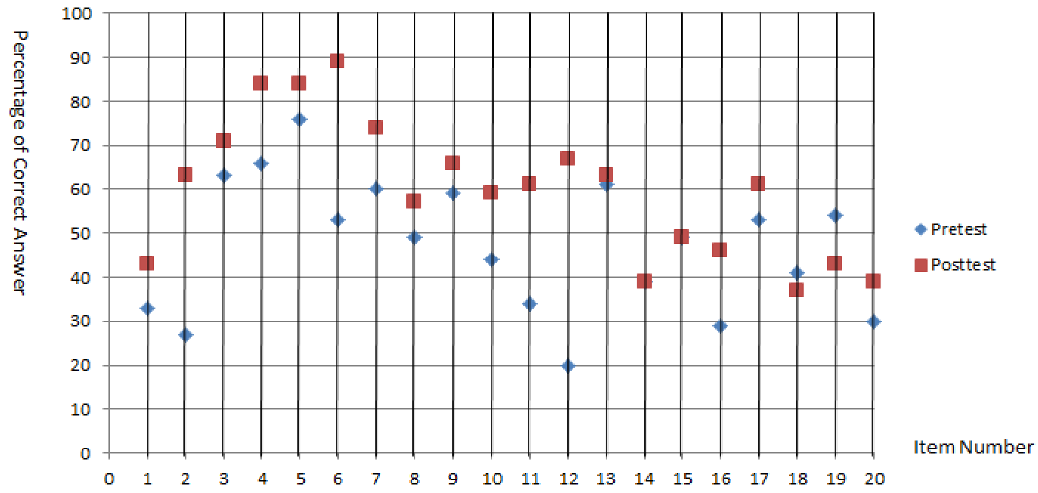



*Figure 4. 2* Comparison of pre-test and post-test scores regarding electricity knowledge

Participants' performance increased from pre-test to post-test in 16 items out of 20. On the other hand, participants' performance stayed same in items 14 and items 15. On the other hand, participants' performance decreased in item 18 and item 19 regarding electricity. Item 14 was an application question and students were expected to connect electricity system of a kettle and the factors affecting wire resistance. This was the most difficult question in the ECK test (P=0.24). Results about this item shows that students still had difficulty in connecting related objective (i.e., factors affecting wire resistance) to their daily life applications (mechanism of kettle) after they took ABI treatment. Likewise, students' performance did not change in item 15. This was an analysis level item and this item aimed to assess participants' identifying variables SPS skills in electricity topic. An experimental design was provided to participants in this item and independent variable was asked to students. Although participants' reported their variables in all their experiments throughout the experiment, only half of the students correctly identified independent variable asked in this question both in pre-test and post-test. Therefore, this result is surprising for item 15. Similarly, there was a slight difference between pre-test and post-test in item 18 in favour of pre-test results. This was a synthesis question examining participants' designing experiment SPS in electricity topic. In this item, students were asked to select an experimental design that seeks the relation between cross sectional area of wire and bulb brightness. Only 40 percent of students selected correct experimental design in pre-test and this percentage was less in post-test. On the other hand, students correctly designed this experiment in last week of the treatment. Therefore, this finding is not expected. Likewise, participants' performance decreased in item 19 which is evaluation level question about formulating hypothesis SPS in electricity topic. In this question, an experimental design was provided to students. Moreover, a hypothesis was presented to students. In this hypothesis, it is said that there is a reverse relationship between cross sectional area and bulb brightness. Moreover, result of the experiment is given in question and hypothesis was falsified according to result. In this item, students were asked to evaluate whether hypothesis is falsified or retained. Moreover, suitable justification was needed to be selected in this question. While more than half of the students correctly answered this question in pre-tests, less than half of them correctly answered this question in post-test. Actually, this was a difficult question because it requires students to know what is hypothesis, know content knowledge about the factors affecting bulb brightness and make reasoning to explain why hypothesis should be rejected or retain. Although this was a difficult question, this does not explain why students' performance decreased in this item. On the other hand, students' performance improved in 16 items out of 20. Accordingly, 4 of them were in knowledge level, 2 of them were in comprehension level, 2 of them in application level, 4 of them were in analysis level, 3 of them were in synthesis level, and 1 of them were in evaluation level. Regarding item content, student improved their conceptual understanding in 10 questions out of 11 and they improved their science process skills in 6 questions out of 9 questions.

This part presented the results regarding the treatment effect on content knowledge; changes on students' epistemological beliefs are addressed in next part.

### 4.2. Results for Changes in Epistemological Beliefs

Second research question of the study was "Is there a change in 6$^{th}$ grade students' epistemological beliefs from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)? " Epistemological belief questionnaire was used as pre-test and post-test to collect data about students' epistemological beliefs. Repeated measure multivariate analysis of variance (MANOVA) was used to identify possible changes in epistemological beliefs components as response to argument based-inquiry treatment. Time was independent variable (i.e. Time1 and Time 2). At Time 1 protest scores were obtained while at Time 2 post test scores were obtained. On the other hand, three components of epistemological beliefs, namely combination of source and certainty (i.e., combsc), development and justification were dependent variables. Prior to repeated measures MANOVA, its assumptions are tested to see whether data fits to conduct repeated measure MANOVA, but first, descriptive statistics about students' epistemological beliefs are presented.

### 4.2.1. Descriptive Statistics for Changes in Epistemological Beliefs.

Table 4.8 presents participants' epistemological beliefs' descriptive statistics for pre-test and post-test.

Table 4. 8

*Descriptive Statistics for Epistemological Beliefs*

| | Pretest | | Posttest | |
|---|---|---|---|---|
| Component | M | SD | M | SD |
| Combination of Source and Certanity | 3.92 | 0.61 | 4.00 | 0.79 |
| Justification | 4.00 | 0.62 | 3.89 | 0.87 |
| Development | 3.67 | 0.59 | 3.74 | 0.80 |

Participants' epistemological beliefs were assessed by EBQ (Conley et al., 2004). Accordingly, there are three components of epistemological beliefs (Özkan, 2008). In line with this, mean and standard deviation of these three components are presented as descriptive statistics for pre-test and post-test. Descriptive statistics suggested that participants marked high scores (e.g. 4-5 over 5) both in pre-tests and post-tests. Accordingly, combination of source and certainty mean score was 3.92 (SD=0.61) in pre-test, and this score slightly increased 4.00 (SD=0.79) in post-test. Likewise, development component was 3.67 (SD=0.57) in pre-test and it reached to

3.74 (SD=0.80) in post-test. Although there were slight increases in these two components from pre-test to post-test, participants' justification component's mean score slightly decreased from 4.00 (SD=0.62) to 3.89 (SD=0.87) from pre-test to post-test. According to descriptive statistics, it can be concluded that participants' justification mean score was highest and combination of source and certainty and development mean scores followed justification in pre-test. On the other hand, participants' combination of source and certainty mean score was highest and justification and development components' mean scores followed combination of source and certainty in post-test.

### 4.2.2. Evaluation of the Assumptions of Repeated Measures Multivariate Analysis of Variance for Epistemological Beliefs.

#### *4.2.2.1. Missing Data Analysis.*

A total of 71 students participated in the study. On the other hand, 66 students completed epistemological beliefs questionnaire in both pre-tests and post-tests. Moreover, one student was found as outlier (as explained in univariate outlier assumption). Therefore, sample size was 65. Accordingly, the analysis was conducted using the data only from these students. This sample size was considered to be sufficient to get valid results. Indeed, according to Hahs-Vaughn (2017), while determining the minimum sample size needed to conduct repeated measures MANOVA, the number of measurement occasions and the number of dependent variables should be considered and the sample size should be greater than the product of these two measures. In the present study, the number of measurement occasions was 2 and the number of dependent variables was 3. Accordingly, the minimum sample size required is 6. Moreover, the Hahs-Vaughn (2017) added that there are some other suggestions for the minimum sample size for MANOVA. For example, according to a recommendation the sample size for MANOVA should be at least 20. When these different suggestions are considered, the sample size of 65 is found to be sufficient to conduct the analysis.

#### *4.2.2.2. Univariate Normality.*

Descriptive statistics were used to test univariate normality assumptions. Skewness and kurtosis values are considered to assess normality. Descriptive statistics for epistemological beliefs components are presented in Table 4.9;

Table 4. 9

*Descriptive Statistics for Epistemological Beliefs Components*

| | Skewness | Kurtosis |
|---|---|---|
| Pre-test | | |
| Combsc* | -0.54 | -0.18 |
| Justification | -0.88 | 0.56 |
| Development | -0.23 | -0.52 |
| | | |
| Post-test | | |
| Combsc* | -1.03 | 0.95 |
| Justification | -1.33 | 1.46 |
| Development | -0.97 | 1.39 |

*Combsc refers the component which is combination of source and certainty.

According to Hahs-Vaughn (2017), distributions having skewness values larger than 2 and kurtosis values beyond 7 are considered as non-normal. On the other hand, Skewness values for epistemological beliefs components are between -1.33 and -0.23. These values are acceptable for normality because they are between acceptable ranges. Likewise, kurtosis values support normality and their range is between -0.52 and 1.46. There is no skewness and kurtosis value that extend critical values.

#### *4.2.2.3. Univariate Outliers.*

Univariate outliers were detected by use of standardized scores. According to Tabachnick and Fidell (2012) standardized scores (z scores) should be between -3.29 and +3.29. The values exceeding these critical values are univariate outliers. Highest and lowest standardized scores for each epistemological belief components are calculated and presented in Table 4.10;

Table 4. 10

*Standardized Scores for Epistemological Beliefs Components*

| | Highest Z-scores | Lowest Z-scores |
|---|---|---|
| Pre-test | | |
| Combsc | 1.76 | -2.59 |
| Justification | 1.62 | -3.07 |
| Development | 2.27 | -2.27 |
| Post-test | | |
| Combsc | 1.26 | -2.68 |
| Justification | 1.26 | -3.31* |
| Development | 1.57 | -3.42* |

*the score is out of acceptable range

According to Table 4.10, one student's standardized score is below the acceptable range regarding post-test justification (-3.31<-3.29). Likewise, same student's post-development score is less than minimum limits for standardized score (-3.42<-3.29). Then, this student's post-test justification score and post-test development score was compared with mean scores for post-test justification and post-test development. While means of post-test justification and post-test development were 3.89 and 3.74 respectively, this student's corresponding scores were 1.00 and 1.00. There were big differences between this student's scores and mean values and so this student's scores were removed from analysis in order not to affect analysis negatively. Moreover, when Mahalanobis Distance was calculated for multivariate outlier, this students' Mahalanobis Distance was found as 22.97 indicating the presence as a multivariate outlier. After this student was removed from analysis, skewness and kurtosis values for each score became closer to zero supporting univariate normality. In conclusion, this outlier was removed prior to analysis.

#### *4.2.2.4. Multivariate Normality for Dependent Variables.*

Linear combination of dependent variables should be normally distributed (Hahs-Vaughn, 2017) for this assumption. Moreover, univariate normality assumption should be met for multivariate normality for dependent variables. In line with this, univariate normality was met and one outlier (Z= -3,31 and -3,42) violating univariate normality was removed from analysis. Removal of this outlier made contribution on multivariate normality for dependent variables.

Moreover, Hahs-Vaughn (2017) claimed that MANOVA including repeated measure MANOVA is robust to violations of multivariate normality when there are 20 or more cases per cell. In this analysis regarding epistemological beliefs, 65 students were included. Hence, it is thought that even though multivariate normality for dependent variable was violated in this study, this does not result in serious conclusions.

Another important point for multivariate normality is multivariate outliers because multivariate outliers might threat multivariate normality (Hahs-Vaughn, 2017). In line with this, next part informs about multivariate outliers.

#### *4.2.2.5. Multivariate Outliers.*

Mahalanobis distance and corresponding probability levels (Tabachnick & Fidell, 2012) were used to decide whether there are multivariate outliers in data. Accordingly, chi square $\chi^2$ table is compared with Mahalanobis distance to claim whether there are multivariate outliers. Degree of freedom (i.e., df) is number of dependent variables and there are three dependent variables in this test. Corresponding chi square value for df =3 is 16.27 and Mahalanobis distance larger than 16.27 is assumed as multivariate outliers (Tabachnick & Fidell, 2012). Mahalanobis distances were calculated from regression analysis using SPSS 26 program. Furthermore, probability values corresponding to Mahalanobis distance are used to determine multivariate normality. Probability levels which are less than 0.001 are evidence that the obtained score is multivariate outlier. Probability levels were computed by the formula 1-CDF.CHISQ (MAH_1, 3) using SPSS. Table 4.11 shows the highest 10 scores for Mahalanobis distance and corresponding probability values;

Table 4. 11

*Highest 10 Scores for Mahalanobis Distance*

| Highest Scores | Mahalanobis Distance | Probability level |
|---|---|---|
| 1 | 22.97* | 0.000* |
| 2 | 10.48 | 0.014 |
| 3 | 10.31 | 0.016 |
| 4 | 9.63 | 0.021 |
| 5 | 8.70 | 0.033 |
| 6 | 8.61 | 0.035 |
| 7 | 8.51 | 0.036 |
| 8 | 7.43 | 0.059 |
| 9 | 7.13 | 0.067 |
| 10 | 6.85 | 0.076 |

According to Table 4.11 the highest score is multivariate outlier because this value's Mahalanobis values exceeds the critical chi square $\chi^2$ value (22.97> 16.27). Likewise, this value's probability level is less than critical value ($p<0.001$). Therefore, it can be concluded that there is a multivariate outlier in data. When data set was examined, this multivariate outlier was the student who was also univariate outlier. Hence, this student's scores were removed from the analysis. Other scores did not threat multivariate outlier assumption.

#### ***4.2.2.6. Linearity.***

Linearity is another assumption for repeated measure MANOVA. Accordingly, there should be straight line relationship between each pair of dependent variables. Scatter plots are used to meet the linearity assumption. Scatter plots used for linearity assumption are given in Appendix O. Matrix scatter plot shows that there is no serious violation of the assumption.

#### ***4.2.2.7. Absence of Multicollinearity and Singularity.***

According to Tabachnick and Fidell (2012), multicollinearity and singularity violations occur when pairs of dependent variables are highly correlated. If the correlation between two dependent variable is higher than 0.90, this is evidence for multicollinearity. Correlation coefficient values were calculated to decide whether there is multicollinearity. Relationships are presented in Table 4.12;

Table 4. 12

*Corelation Coefficient Values for Epistemological Beliefs Components*

| | Pre-test combsc | Post-test combsc | Pre-test justification | Post-test justification | Pre-test development | Post-test development |
|---|---|---|---|---|---|---|
| Pre-test combsc | | 0.49 | 0.56 | 0.09 | 0.27 | 0.03 |
| Post-test combsc | 0.49 | | 0.31 | 0.55 | 0.14 | 0.45 |
| Pre-test justification | 0.56 | 0.31 | | 0.40 | 0.31 | 0.21 |
| Post-test justification | 0.09 | 0.55 | 0.40 | | 0.23 | 0.84 |
| Pre-test development | 0.27 | 0.14 | 0.31 | 0.23 | | 0.20 |
| Post-test development | 0.03 | 0.45 | 0.21 | 0.84 | 0.20 | |

According to Table 4.12, there are no correlation between variables higher than 0.90. Therefore, it can be claimed that there is no multicollinearity in data. Singularity is another concern and singularity happens when variables are combination of other variables. In this study, variables were obtained from Özkan (2008). Confirmatory factor analysis provided three structures in this previous study which are combination of source and certainty, justification and development. Therefore, it is assumed that none of the epistemological belief component is combination of others and there is no singularity violation in data.

In conclusion, data obtained from epistemological belief questionnaire has no serious violations for repeated measure MANOVA assumptions. Therefore, repeated measure MANOVA test can be conducted to reveal whether there is a change on epistemological beliefs as response to ABI treatment.

### 4.2.3 Repeated Measures of Multivariate Analysis of Variance for Epistemological Beliefs.

Repeated measure MANOVA was conducted to see the impact of ABI treatment on students' epistemological beliefs including three components namely combination of source and certainty, justification and development. Results of Repeated measure MANOVA were given in Table 4.13. According to the results there was no significant effect of argument-based inquiry treatment on combined dependent variables, $F(3,62)= 2.24$, $p= 0.09$; Wilks' Lambda= 0.90. Due to fact that treatment had no effect on combined variables, separate analysis for dependent variables were not conducted.

Table 4. 13

*Multivariate Test Result within Subjects for Epistemological Beliefs*

| Effect | Wilks'Lambda | F | Hypothesis df | Error df | Significance Level |
|---|---|---|---|---|---|
| Time | 0.90 | 2.24 | 3 | 62 | 0.09 |

In order to get more information about changes on participants' epistemological beliefs, participants' level of agreement to each epistemological beliefs questionnaire items were taken into consideration for both pre-test and post-test. Accordingly, percentages of strongly agree and agree were summed under Agree title. Likewise, percentages of strongly disagree and disagree were summed under Disagree title. Moreover, percentages of the undecided response for each item are recorded too. Table 4.14 presents combination of source and certainty items and percentages of students' answers including agreement, disagreement and undecided answers for pre-test and post-test. Source and certainty items were negative items (Conley et al., 2004) therefore; disagreement with related items suggests sophisticated epistemological beliefs.

Table 4. 14

*Percentages of Students' Response to Combination of Source and Certainty Items*

| Item | Pretest Disagree (%) | Pretest Undecided (%) | Pretest Agree (%) | Posttest Disagree (%) | Posttest Undecided (%) | Posttest Agree (%) |
|---|---|---|---|---|---|---|
| Everybody has to believe what scientists say. | **84.8** | 7.6 | 7.6 | **83.3** | 6.1 | 10.6 |
| In science, you have to believe what the science books say about stuff. | **84.3** | 9.1 | 4.5 | **81.8** | 7.6 | 10.6 |
| Whatever the teacher says in science class is true. | **63.6** | 18.2 | 18.2 | **74.2** | 12.2 | 13.6 |
| If you read something in a science book, you can be sure it is true. | **51.5** | 30.3 | 18.2 | **57.6** | 33.3 | 9.1 |
| Only scientists know for sure what is true in science. | **79.8** | 13.6 | 7.6 | **77.3** | 15.2 | 7.5 |
| Scientists pretty much know everything about science; there is not much more to know. | **80.3** | 6.1 | 13.6 | **69.7** | 19.7 | 10.6 |
| Scientific knowledge is always true. | **65.2** | 22.7 | 12.1 | **59.1** | 28.8 | 12.1 |
| Once scientists have a result from an experiment, that is the only answer. | **78.8** | 9.1 | 12.1 | **75.8** | 12.1 | 12.1 |
| Scientists always agree about what is true in science. | **54.5** | 19.7 | 25.8 | **63.6** | 21.2 | 15.4 |

According to Table 4.14, students' level of disagreement was higher than their agreement and undecided response on all negative items found in combination of source and certainty component for both pre-test and post-test. This shows that students marked the options which are consistent with sophisticated epistemological beliefs regarding this component because all the items in this component were negative item.

In pre-test, the highest percentage of disagreement was 84.8 in item “Everybody has to believe what scientists say.” and second highest percentage of disagreement was 84.3 in item “In science, you have to believe what the science books say about stuff.” On the other hand, the lowest percentage of disagreement was 51.5 in item that is “If you read something in a science book, you can be sure it is true.” Likewise, students’ level of disagreement was relatively low (54.5 %) in the item “Scientists always agree about what is true in science.”

Similarly, the highest percentage of disagreement was 83.3 in the same item “Everybody has to believe what scientists say.” in post-test. Likewise, the item “In science, you have to believe what the science books say about stuff.” followed it with 81.8 %. Furthermore, the lowest percentage of disagreement (57.6 %) was still the item “If you read something in a science book, you can be sure it is true.”

When change of students’ disagreement level was examined from pre-test to post-test, it is seen that percentages of disagreements for three items increased from pre-test to post-test. Accordingly, disagreement increased from 63.6 % to 74.2 % in item “Whatever the teacher says in science class is true.”, from 51.5 % to 57.6 % in item “If you read something in a science book, you can be sure it is true.” and from 54.5 to 63.6 in item “Scientists always agree about what is true in science.” Although percentages of disagreement increased in these three items, it decreased obviously in two items. Accordingly, students disagreement percentage decreased from 80.3 to 69.7 in item “Scientists pretty much know everything about science; there is not much more to know.” and decreased from 65.2 to 59.1 in item “Scientific knowledge is always true.” In conclusion, the data suggests that students had sophisticated epistemological beliefs in source and certainty components’ combination both in pre-test and post-test because disagreement was higher than agreement and undecided response for all items written in negative format. Moreover, their percentage did not change too much according to percentages of responses obtained in pre-test and post-test. On the other hand, when three items in which disagreement level increased are examined, it might be said that students became a bit more sceptical to source of knowledge and students started to think that scientists may disagree with each other. On the other hand; after this study, students' approval to some positivist epistemology might increase because the percentages of ideas which are scientists know everything in science and scientific knowledge is always true increased from pre-test to post-test. However, these claims are still not eligible because students' responses to items in general seem to be stable.

After combination of source and certainty component was reported based on percentages of each item, justification item's percentages are reported. Table 4.15 presents the percentages of each item related with justification component of epistemological beliefs.

Table 4. 15

*Percentages of Students' Response to Justification Items*

| Item | Pretest Agree (%) | Pretest Undecided (%) | Pretest Disagree (%) | Posttest Agree (%) | Posttest Undecided (%) | Posttest Disagree (%) |
|---|---|---|---|---|---|---|
| Ideas about science experiments come from being curious and thinking about how things work. | **72.7** | 22.7 | 4.6 | **63.6** | 21.2 | 15.2 |
| In science, there can be more than one way for scientists to test their ideas. | **84.8** | 9.1 | 6.1 | **72.7** | 13.6 | 13.7 |
| One important part of science is doing experiments to come up with new ideas about how things work. | **77.3** | 9.1 | 13.6 | **75.8** | 9.1 | 15.1 |
| It is good to try experiments more than once to make sure of your findings. | **74.2** | 15.2 | 10.6 | **74.2** | 15.2 | 10.6 |
| Good ideas in science can come from anybody, not just from scientists. | **75.8** | 9.1 | 15.1 | **72.7** | 10.6 | 16.7 |
| A good way to know if something is true is to do an experiment. | **81.8** | 6.1 | 12.1 | **71.2** | 13.6 | 15.2 |
| Good answers are based on evidence from many different experiments. | **66.7** | 27.3 | 6.0 | **68.2** | 24.2 | 7.6 |
| Ideas in science can come from your own questions and experiments. | **69.7** | 15.2 | 15.1 | **69.7** | 15.1 | 15.2 |
| It is good to have an idea before you start an experiment. | **81.8** | 6.1 | 12.1 | **71.8** | 9.1 | 19.7 |

As it was reported in combination of source and certainty, students had sophisticated epistemological beliefs regarding justification component because their agreement was higher than undecided response and disagreement for all items both in pre-test and post-test.

In pre-test, the highest approval was 84.8 % in the item “In science, there can be more than one way for scientists to test their ideas.” and there were two items getting second highest approval which are “A good way to know if something is true is to do an experiment.” and “It is good to have an idea before you start an experiment.” These two items’ approval percentage was 81.8 %. On the other hand, the least approved item was “Good answers are based on evidence from many different experiments.”. Accordingly, 66.7 % of students agreed to this item.

On the other hand, it can be said there is a decrease in students’ agreement with justification items in post-test. According to post-test items, the highest agreement was with  the item “One important part of science is doing experiments to come up with new ideas” (75.8 % ), the second highest agreement in post test was with the item “It is good to try experiments more than once to make sure of your findings.” (74.2) However, the least agreed item was “Ideas about science experiments come from being curious and thinking about how things work.”. Accordingly, 63.6 % of students agreed with this item in post-test.

Although students’ agreement percentages with justification items were higher than disagreement and undecided response’s percentages both in pre-test and post-test for all items, there is no obvious increase on students’ agreement to justification items from pre-test to post-test in item level. On the other hand, students’ agrement with justification items apparently decreased in four items. The items that students’ agreement decreased from pre-test to post-test are “Ideas about science experiments come from being curious and thinking about how things work.”, “In science, there can be more than one way for scientists to test their ideas.”, “A good way to know if something is true is to do an experiment.” and “It is good to have an idea before you start an experiment.” All these four items’ agreement decreased around 10 % in average. In conclusion, it can be said that participants’ epistemological beliefs were sophisticated both in pre-test and post-test in terms of justification component. However, these beliefs started to slightly decrease after treatment because related percentages showed that students' agreement with several justification items decreased from pre-test to post-test. In spite of this decrease, high level of agreement with the items suggested that students' epistemological beliefs regarding justification dimension reflected sophisticated views both before and after the treatment

Finally, percentages of students' responses about development component are presented. These percentages further inform about students' epistemological beliefs in development component in item level. Percentages are presented in Table 4.16.

Table 4. 16

*Percentages of Students Response to Development Items*

| | Pretest | | | Posttest | | |
|---|---|---|---|---|---|---|
| Item | Agree (%) | Undecided (%) | Disagree (%) | Agree (%) | Undecided (%) | Disagree (%) |
| Some ideas in science today are different than what scientists used to think. | **45.5** | 37.9 | 16.6 | **50.0** | 40.9 | 9.1 |
| The ideas in science books sometimes change. | **59.1** | 33.3 | 7.6 | **56.1** | 30.3 | 13.6 |
| There are some questions that even scientists cannot answer. | **80.3** | 10.6 | 9.1 | **72.7** | 7.6 | 19.7 |
| Ideas in science sometimes change. | **72.7** | 13.6 | 13.7 | **77.3** | 12.1 | 10.6 |
| New discoveries can change what scientists think is true. | 42.4 | **47.0** | 10.6 | **47.0** | 42.4 | 10.6 |
| Sometimes scientists change their minds about what is true in science. | **62.1** | 21.2 | 16.7 | **66.7** | 25.8 | 7.5 |

According to Table 4.16, participants percentages of agreement with development items were higher than their percentages of undecided response and disagreement in general both in pre-test and post-test. Therefore, it can be said that students had sophisticated epistemological beliefs regarding development component both in pre-test and post-test. However, participants' agreement level with development items is lower comparing with their agreement to other two epistemological beliefs components (e.g. justification).

Percentages obtained from pre-test shows that agreement percent is higher than disagreement and undecided percent in five items out of six. On the other hand, undecided response percentage was higher than agreement percentage in one item (47.0 % and 42.4 %) which is “New discoveries can change what scientists think is true.” The highest percentage that students agreed in pre-test was 80.3 in the item “There are some questions that even scientists cannot answer.” and the item “Ideas in science sometimes change.” followed this item with 72.7 %. On the other hand, the item that students least agreed was “New discoveries can change what scientists think is true.” with 42.4 %.

While percentages of students' agreement were higher than undecided response and disagreement percentage in five items in pre-tests, agreement percentage was higher than other responses' percentages in all development items in post-tests. Students agreed most with the item “Ideas in science sometimes change.” (77.3 %). The item “There are some questions that even scientists cannot answer” followed the most agreed item with an agreement of 72.7 % in post-test. On the other hand, the least agreed item was “New discoveries can change what scientists think is true.” (47.0 %) in post-test.

Table 4.16 also informs about changes about students' beliefs in development component of epistemological beliefs from pre-test to post-test. Accordingly, percentages of students' agreement increased in four items and the increase percent for these items were around 4-5 %. For example, students' agreement increased from 45.5 % to 50.0 % in item “Some ideas in science today are different than what scientists used to think.” On the other hand, percentages of students' agreement clearly decreased in one item "There are some questions that even scientists cannot answer." from 80.3 to 72.7. Likewise, students' agreement with the item "The ideas in science books sometimes change." slightly decreased from pre-test to post-test. In conclusion, it can be claimed that students had sophisticated epistemological beliefs about development both in pre-test and post-test although there are some increase or decrease on percentages of students' agreement in some items from pre-test to post-test.

In conclusion, percentage tables were presented to portray how students' epistemological beliefs changed according to each component in item level. These three tables showed that students held sophisticated epistemological beliefs in both pre-test and post-test scores because percentages of students' agreement were almost more than the percentages of undecided response and disagreement. Although there were slight differences between pre-test and post-test in item level, general trend remained stable for all three components. These results obtained from percentage tables is also consistent with the inferential statistics' (i.e. Repeated MANOVA) results revealing that there is no statistical change on students' epistemological beliefs from pre-test to post-test. The following part continues with result about change in participants' science process skills as response to treatment.

### 4.3. Results for Changes in the Science Process Skills

Third research question of the study was "Is there a change in 6$^{th}$ grade students' science process skills from Time 1 (before the ABI treatment) to Time 2 (after the ABI treatment)?" Science process skills test developed by Burns et al. (1985) was used to collect data regarding students' science process skills. This test was developed to assess integrated process skills and the test was non-specific to curriculum. Therefore, results of this test reflect students integrated process skills (e.g., operational definition) about general science domain.

Prior to analysis, within subject repeated measure MANOVA was planned to be conducted because there were five different components of science process skill test. However, these components had only few items and these components' reliability scores were too low. Therefore, SPS was decided to be examined as a unidimensional construct rather than combination of five dimensions because of low reliability scores of dimensions. Similar approach (e.g. accepting SPS as one dimension) was also used in previous research (Cansız, 2014; Gök, 2014). Hence, within subject repeated MANOVA was not conducted to reveal whether corresponding components changed. Lack of separate analysis of science process skills components can be one of the limitations of current study. On the other hand, paired sample t-test was conducted on total instrument scores which is the combination of other components. This part includes assumptions of paired sample t-test and its results. First, assumptions of the test are explained and then results of paired sample-t test are presented. Prior to evaluation of assumptions of pair sample t-test, descriptive statistics for science process skill scores obtained in pre-test and post-test are presented.

### 4.3.1. Descriptive Statistics for Changes in Science Process Skills.

TIPSII included 26 multiple choice test items (Can, 2008). Thus, if all questions are answered correctly, the maximum score will be 26. Table 4.17 presents descriptive statistics for science process skills including sample size, mean, standard deviations, minimum, maximum values and range.

Table 4. 17

*Science Process Skills Descriptive Statistics*

| | N | Mean | Standard Deviation | Minimum | Maximum | Range |
|---|---|---|---|---|---|---|
| Pre-test | 63 | 14.70 | 5.48 | 5 | 25 | 20 |
| Post-test | 63 | 14.33 | 5.98 | 4 | 25 | 21 |

Descriptive statistics show that mean scores of students’ science process skills were similar in pre-test and post-test. The mean score of pre-test was 14.70 (SD=5.48) and this score slightly decreased to 14.33 (SD=5.98) in post-test. The minimum SPS score was 5 and the maximum score was 25 over 26 in pre-test. Similarly, the minimum SPS score was 4 and the maximum score was 25 in post-test. The range also did not change from pre-test to post-test.

### 4.3.2. Evaluation of the Assumptions of Paired Sample t-test for Science Process Skills.

There are two basic assumptions of paired sample t-test namely, independence of observations and normal distribution of difference scores in the population (Gravetter & Wallnau, 2007). Independence of observation assumption requires that observation is independent between cases, not across occasion or time. In the current study, while administering the instruments, no interaction among the students was allowed. There was no cheating or discussions among the students. Thus, this assumption is assumed to be satisfied. Regarding normality assumption, the difference scores obtained from subtraction of pre-test scores from post-test scores were calculated and skewness and kurtosis values were examined for the difference scores (see Table 4.18).

Table 4. 18

*Descriptive Statistics for Difference Scores*

| | Mean Difference | Standard Deviation | Skewness | Kurtosis | Minimum | Maximum |
|---|---|---|---|---|---|---|
| The difference between post-test and pre-test | -0.36 | 4.91 | -0.12 | -0.03 | -13 | 10 |

According to Table 4.18; skewness values for difference is -0.12 and kurtosis value is -0.03 and these values are between acceptable ranges. Therefore, these results do not violate normality assumption.

### 4.3.3. Paired Sample t-test for Science Process Skills.

Paired sample t-test was conducted to see the effect of argument based inquiry treatment on students' integrated science process skills. Results of paired sample t-test were given in Table 4.19;

Table 4. 19

*Result of Pair Sample t-test regarding Science Process Skills*

| Pair | T | Df | Sig. | Mean Difference | Confidence Interval (95 %) | |
|---|---|---|---|---|---|---|
| | | | | | Lower | Upper |
| Difference between post-test and pre-test | -0.59 | 62 | .557 | -0.36 | -1.60 | 0.87 |

According to the Table 19, there was no statistically significant change from pre-test (M=14.70, SD=5.48) to post-test (M=14.33, SD=5.98), t (62) = -0.59, p>0.05 (two tailed). The mean difference score was -0.36 for science process skills scores with 95 % confidence interval ranging -1.60 to 0.87. It can be claimed that the treatment did not have a significant effect on students' integrated science process skills.

In order to better portray changes in science process skills, results were examined in item level. Figure 4.3 presents changes of percentage of correct answers in science process skill test from pre-test to post-test for each item

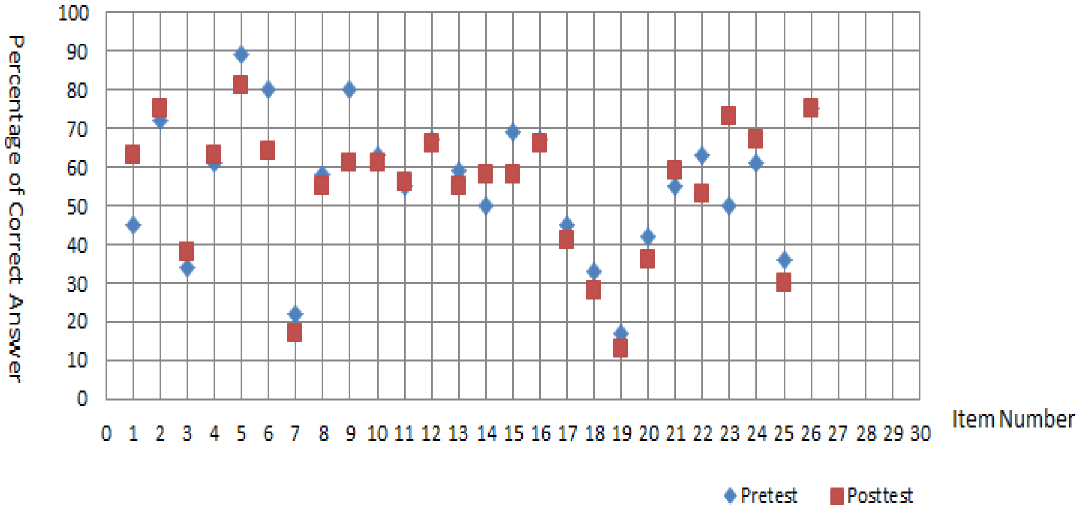



*Figure 4. 3* Comparison of pre-test and post-test scores regarding science process skills

According to Figure 4.3, there is no specific trend for changes of percentages of correct answers from pre-test to post-test. Percentages of correct answers were similar in 18 items out of 26 because blue and red squares were close to each other in these items. In these items (e.g. item 2, item 7), differences of percentages of correct answers were less than 10 %. On the other hand, percentages of correct answer were clearly higher in four items in pre-test scores compared with post-test scores (item 6, item 9, item 15, item 22). Content of these items were different from each other. For example, item 6 was about graphing, and item 9, 15 and 22 were about identifying variables (e.g. control variable) integrated process skills. On the other hand, percentages of correct answers increased in four items which are item 1, item 14, item 23 and item 24. Content of item 1 was about defining operationally, and content of item 14, 23, and 24 were about identifying variables.

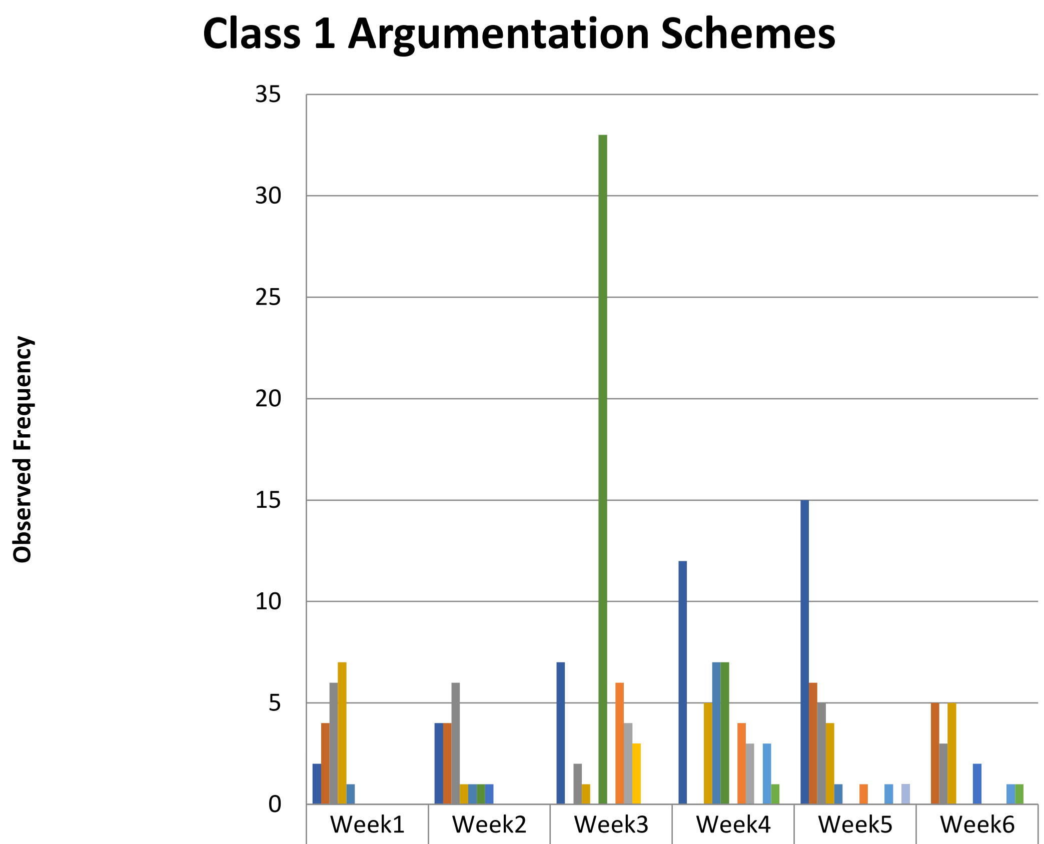
To sum up, it can be claimed that percentages of correct anwers regarding science process skills was stable in general because the scores obtained in pre-test and post-test were close to each other in 18 items out of 26 (The percentage of difference was less than 10 % in these 18 items). Moreover, the numbers of items that percentages of correct answers decreased or increased in this study are close each other (percentage of correct answer decreased in 4 items and increased in 4 items too. Percentages of difference between pre-test and post-test were more than 10 % in these items), thus an inference cannot be done based on number of items to reach a conclusion (i.e. increase or decrease) for changes. Even though percentage of correct answers obviously increased or decreased from pre-test to post-test in some items, content of items did not diverge so much between items in which percentages of correct answers increased and items in which percentages of correct answers decreased. For example, three of the items in which percentages of correct answers increased were about identifying variables. Likewise, three of the items in which percentages of correct answers decreased were about identifying variables too. Therefore, it cannot be claimed that percentages of correct answers changed based on content of the item (e.g., identifying variables). These percentages of correct answers are also consistent with paired sample t-test results showing there is no statistical difference between pre-test and post-test scores in terms of science process skills. The next part presents the results about argumentation schemes.

### 4.4. Findings for Nature of Students Arguments

Forth research question of the study is "What is the nature of 6$^{th}$ grade students' arguments when they are analyzed based on argumentation schemes?" Walton's (1996) argumentation schemes (see Table 3.6) were used to reveal students' argumentation schemes throughout the process. In addition to Walton's explanations of argumentations schemes, Duschl's (2007) ideas about argumentation schemes were also used as a guide to reveal students' argumentation schemes. Total frequency of each argumentation schemes and their percentages are presented for both classes in each week. Firstly, class 1's argumentation schemes and secondly, class 2's argumentation schemes are presented. After findings of both classes are given, similarities and differences derived from both classes' findings are presented as assertions. These assertions summarize 6$^{th}$ grade level students' argumentation schemes and can be thought as the answer of research question 4.

#### 4.4.1. Class 1 Findings.

Class 1’s whole class discussions were analysed based on argumentation schemes. Students’ argumentation schemes are linked with students’ reasoning about the argument (i.e., product of argumentation). According to Walton (1996), people can use 25 different schemes when they propose an argument (see chapter 3, Table 3.7). Similar to Walton, Duschl (2007) also used some of these argumentation schemes, but Duschl did not add some new argumentation schemes. Therefore, current study focused on all 25 argumentation schemes. Total frequencies of argumentation schemes used by class 1 students for each week are presented in Figure 4.4. Content of the each week, how ABI was applied for each week and type of data used in corresponding week was summarized in chapter 3 Table 3.2.




| | Week1 | Week2 | Week3 | Week4 | Week5 | Week6 |
|---|---|---|---|---|---|---|
| Sign | 2 | 4 | 7 | 12 | 15 | 0 |
| Evidence to Hypothesis | 4 | 4 | 0 | 0 | 6 | 5 |
| Comparison | 6 | 6 | 2 | 0 | 5 | 3 |
| Position to Know | 7 | 1 | 1 | 5 | 4 | 5 |
| Example | 1 | 1 | 0 | 7 | 1 | 0 |
| Expert | 0 | 1 | 33 | 7 | 0 | 0 |
| Popularity | 0 | 1 | 0 | 0 | 0 | 2 |
| Corelation to cause | 0 | 0 | 6 | 4 | 1 | 0 |
| Consequence | 0 | 0 | 4 | 3 | 0 | 0 |
| Committment | 0 | 0 | 3 | 0 | 0 | 0 |
| Cause effect | 0 | 0 | 0 | 3 | 1 | 1 |
| Analogy | 0 | 0 | 0 | 1 | 0 | 1 |
| Gradualism | 0 | 0 | 0 | 0 | 1 | 0 |

*Figure 4. 4* Total frequencies of argumentation schemes used by class 1

Firstly, there are 25 possible argumentation schemes (Walton, 1996). However, Figure 4.4 suggested that class 1 students used 13 argumentation schemes out of 25 schemes (e.g. argument from sign) in whole class discussions and so 12 argumentation schemes (e.g., argument from established rule) were not used by Class 1.

Secondly, students' argumentation schemes can be explained based on total number of argumentation schemes used and total types of argumentation schemes that students used in class. Therefore, Figure 4.4 firstly was interpreted based on total number of argumentation schemes and then this figure was interpreted depending on the types of argumentation schemes in total.

Regarding total number of argumentation schemes, argumentation schemes were mostly used in week 3 in class 1 where students used 56 argumentation schemes. Second week that students used argumentation schemes most was week 4 where students used 42 argumentation schemes. Students were also used argumentation schemes frequently in week 5. Total number of argumentation schemes used in week 5 was 34. On the other hand, Class 1 students did not actively used argumentation schemes in week 1, week 2 and week 6. Accordingly, total number of argumentation schemes used by class 1 students was 20 in week 1, 18 in week 2 and 17 in week 6. In conclusion, the order of the weeks depending on the total number of argumentation schemes is: Week 3> Week 4> Week 5> Week 1> Week 2 > Week 6

Total number of argumentation schemes used in class 1 for each week can also be seen from Figure 4.5 linear line graph:

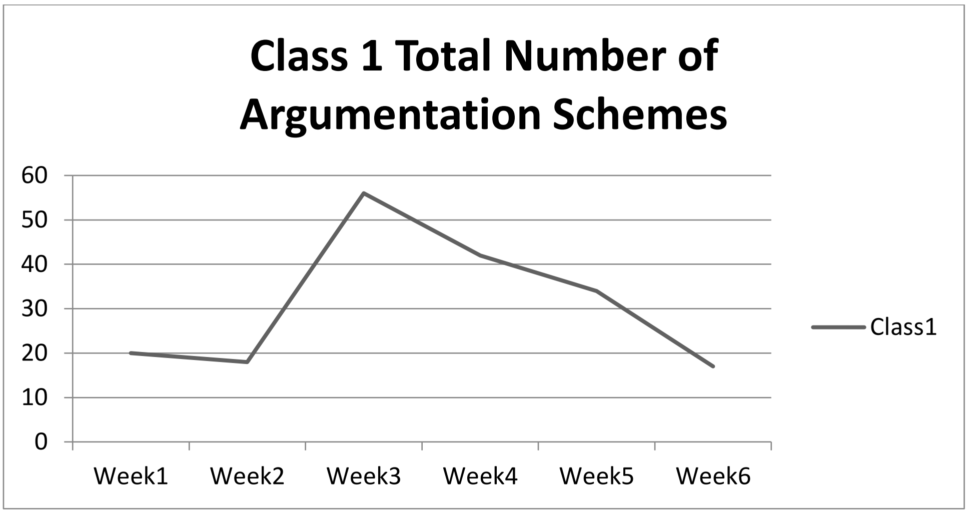



*Figure 4. 5* Class 1 argumentation schemes weekly line graph

Figure 4.4 also informs the types of argumentation schemes that students used in each week. The number of type of argumentation schemes that class 1 students used was most in week 4 and week 5. In these weeks, 8 different types of argumentation schemes were used in class 1. The argumentation schemes used by students in week 4 were sign, position to know, example, expert, corelation to cause, consequence, cause effect and analogy. On the other hand, sign, evidence to hypothesis, comparison, position to know, example, corelation to cause, cause effect, and gradualism were the types of argumentation schemes used in week 5. Besides, 7 different argumentation schemes were used in week 2 and week 3. While sign, evidence to hypothesis, comparison, position to know, example, expert and popularity schemes were used in week 2; sign, comparison, position to know, expert, corelation to cause, consequence, and commitment were used in week 3 regarding the types of argumentation schemes. Class 1 students used 6 different types of argumentation schemes in week 6. The schemes that students used in this week were evidence to hypothesis, comparison, position to know, popularity, cause effect and analogy. The fewest numbers of types of argumentation schemes were used in week 1 in class 1. Class 1 students used sign, evidence to hypothesis, comparison, position to know and example schemes in this week. In conclusion, class 1 students started with using five different types of argumentation schemes in week 1 and number of types of argumentation schemes increased to eight different schemes in last weeks except week 6. Therefore, it can be said that class 1 students increased the use of number of argumentation schemes types over time in the study.

When frequency of use of argumentation schemes are compared with each other, Figure 4.4 shows that the most frequently used scheme was argument from expert opinion which was used 41 times. The second most frequently schema was argument from sign and this scheme was used 40 times. Argument from position to know followed these two schemes. Accordingly, class 1 students used position to know scheme 23 times. Similarly, class 1 used argument from comparison 22 times and argument from evidence to hypothesis 19 times throughout the study. Students also used argument from corelation to cause (11 times), argument from example (10 times), and argument from consequence (7 times) although these schemes were not used as often as the most frequently used schemes. On the other hand, other schemes which were used by class 1 students were argument from cause to effect (5 times), argument from popularity (3 times), argument from commitment (3 times), argument from analogy (2 times), and argument from gradualism (1 time).

While frequency graph (Figure 4.4) informed about students' total number of argumentation schemes and types of argumentation schemes students used, pie charts were prepared to compare percentages of each argumentation schemes with others for each week. These pie charts are presented in Figure 4.6;

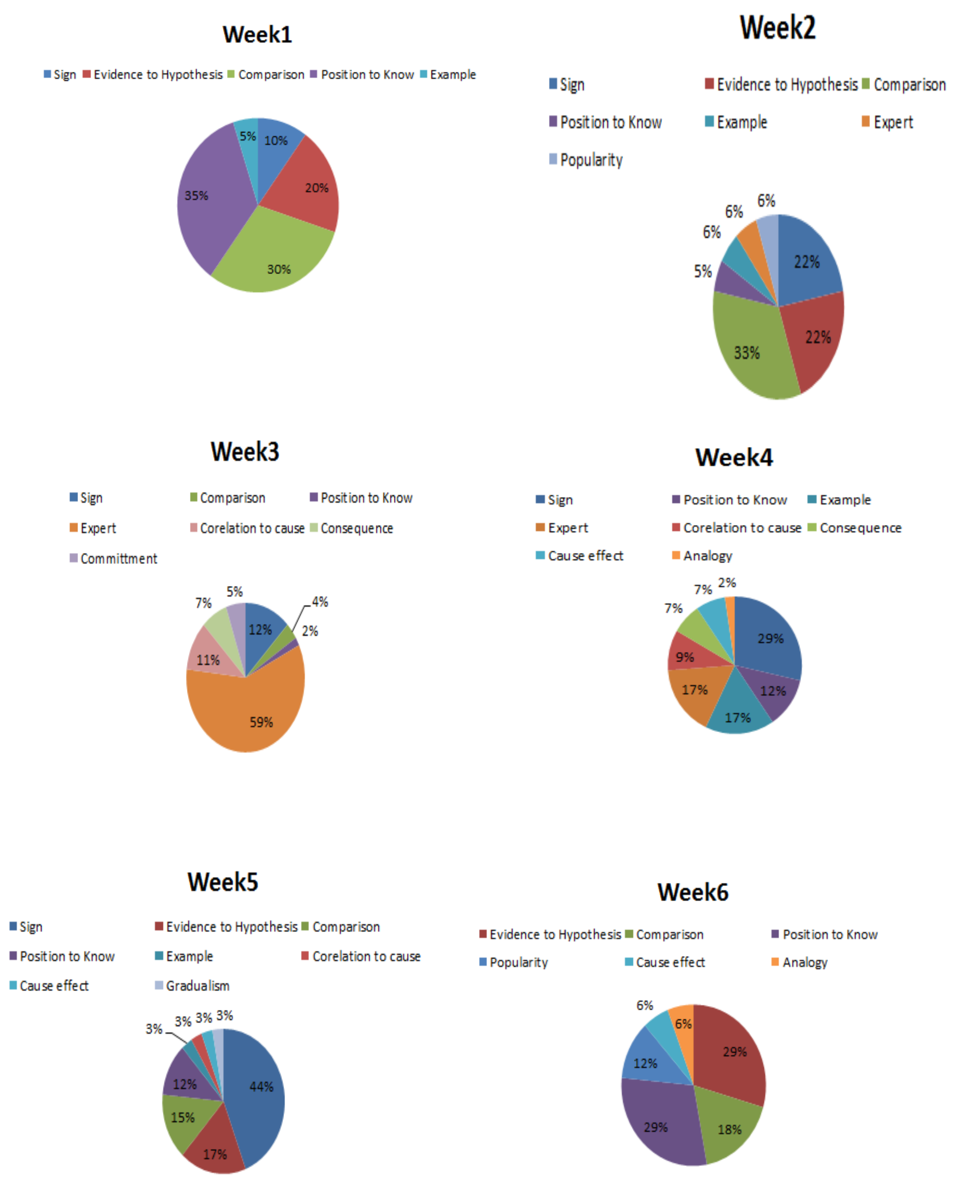



*Figure 4. 6* Class 1 weekly percentages of argumentation schemes

According to Figure 4.6, class 1 students mostly used argument from position to know, argument from comparison, and argument from evidence to hypothesis in week 1 which was about heat conductivity. Percentages of these schemes were 35 %, 30 % and 20 % respectively. Argument from sign and argument from example on the other hand were used less. While 10 % of the arguments were argument from sign in this week, 5 % of arguments were argument from example. After students presented their arguments, other students asked questions to learn more about experiments and results, so students mostly used argument from position to know in this week. While students presented their arguments, they focused on the comparison of concrete materials such as metal spoon and plastic spoon. Therefore, students used argument from comparison very much when they constructed their arguments. Likewise, students formed hypothesis prior to experiment and tested this hypothesis and reached some evidence regarding the validity of hypothesis. By doing this, students frequently used argument from evidence to hypothesis. On the other hand, students could not explain the observed phenomena in some cases and they formulated ideas using their inferences. By this way, students used argument from sign. Moreover, students sometimes referred examples to support their claims and so they referred argument from examples in this week.

Similar to week 1, class 1 students used argument from comparison most in week 2. One third of the arguments were arguments from comparison in this week. Argument from sign and argument from evidence to hypothesis were other two schemes that were frequently used in week 2. Percentages of each of these two argumentation schemes were 22 %. Apart from these three argumentation schemes, students used argument from example, expert, popularity and position to know. However, percentages of these four argumentation schemes were fewer and corresponded to 5 or 6 % of total argumentation schemes. In week 2, the topic was about heat insulation. Students conducted experiments to reveal which materials were better heat insulators. While forming their arguments, they compared different cups such as metal cup and foam cup with each other, so they frequently used arguments from comparison. Likewise, they tested their hypothesis, and so they consulted argument from evidence to hypothesis in this week. Similarly, argument from sign was often used by students when there were uncertainties. Students used inferences to reach explanation and eliminate uncertainties. By using their inferences, they used argument from sign. While students were trying to make explanations using their inferences, they were aware that their explanations were not certain and there were some possibilities that their ideas can be wrong. Although students actively used argument from position to know in week 1, they did not use this scheme frequently in week 2.

In week 3, students did not conduct experiments and they used evidence cards to construct their arguments in thermal insulating products. More than half of the argumentation schemes were argument from expert which corresponded to 59 % of all schemes in this week. Argument from sign (12 %) and argument from correlation to cause (11 %) were other two noticeable schemes used in week 3. Students also used argument from consequence (7 %), argument from commitment (5 %), argument from comparison (4 %) and argument from position to know (2 %). However, percentages of these four schemes were fewer. According to this data, students mainly consulted on evidence cards when they constructed their arguments; therefore, argument from expert opinion dominated this week. When uncertainties emerged they formulated idea to clarify issue by using their inferences. In some points, they knew that their ideas had possibility of being wrong, so their argumentation schemes were argument from sign. However, in some points, students were certain that their ideas were correct even though their ideas were wrong, therefore, it is thought that students used argument from correlation to cause schema in such situations.

Topic of the week 4 was energy sources. Three most common argumentation schemes that were used in this week were argument from sign, argument from example, and argument from expert. While percentage of argument from sign corresponded to 29 % of all arguments, percentages of argument from example and argument from expert were 17 %. Argument from position to know and argument from correlation to cause followed these three schemes. Percentages of these two argumentation schemes were 12 % and 9 % respectively. Three more argumentation schemes which were argument from consequence, argument from cause to effect and argument from analogy were also used in this week, but percentage of these schemes were fewer. Students had prior knowledge in energy sources topic because they actively discussed in both pre-activity discussions and whole class discussion. Moreover, evidence cards supported their argumentation. Likewise, topic was closely related with daily life. During argumentation, students frequently used their inferences and so argument from sign became most used argumentation scheme. Likewise, students consulted on the information given in evidence cards and this supported their use of argument from expert opinion. Furthermore, students frequently used argument from example to support their ideas benefitting from the direct link between daily life experiences and the topic. When students did not have sufficient knowledge, they asked questions to each other. This caused them to use argument from position to know during whole class discussions. On the other hand, students connected two different phenomena as cause and effect actually there was no cause effect relationship between them and this situation resulted in using argument from correlation to cause. However, students sometimes noticed the correct cause effect relationships and they used argument from cause to effect schemes in such instances. Likewise, students sometimes considered possible benefits and damages of using relevant energy sources, so they used argument from consequence scheme.

Students started electricity unit in week 5 and the topic was electrical conductivity. In line with this, students conducted their experiments. Nearly half of the argumentation schemes revealed in class 1 in this week were argument from sign (44 %). Argument from evidence to hypothesis, argument from comparison and argument from position to know followed this scheme. The percentages of these three argumentation schemes were 17 %, 15 % and 12 % respectively. Students also rarely used argument from example, correlation to cause, cause-effect and gradualism in this week. Students tried to explain their observations and their effort was based on making inferences to explain observed phenomena. For example; they tried to explain why salty water conducted electricity, and they used their inferences and so they used argument from sign. Second most used scheme was argument from evidence to hypothesis. Students used this scheme often in constructing their hypothesis prior to testing and they pointed out this scheme when they presented their results in whole class discussions. Regarding use of argument from comparison, students compared concrete materials throughout the discussion. For example, they mentioned the examples of electricity conductors such as metal cup and compared them with materials that do not conduct electricity such as foam cup. By this way, they actively used argument from comparison. Likewise, students examined other groups by asking questions and so they used argument from position to know in this week too. For example, one of the students asked presenting group the relationship between battery power and bulb brightness when presenting groups explained the relation between thickness of wire and bulb brightness.

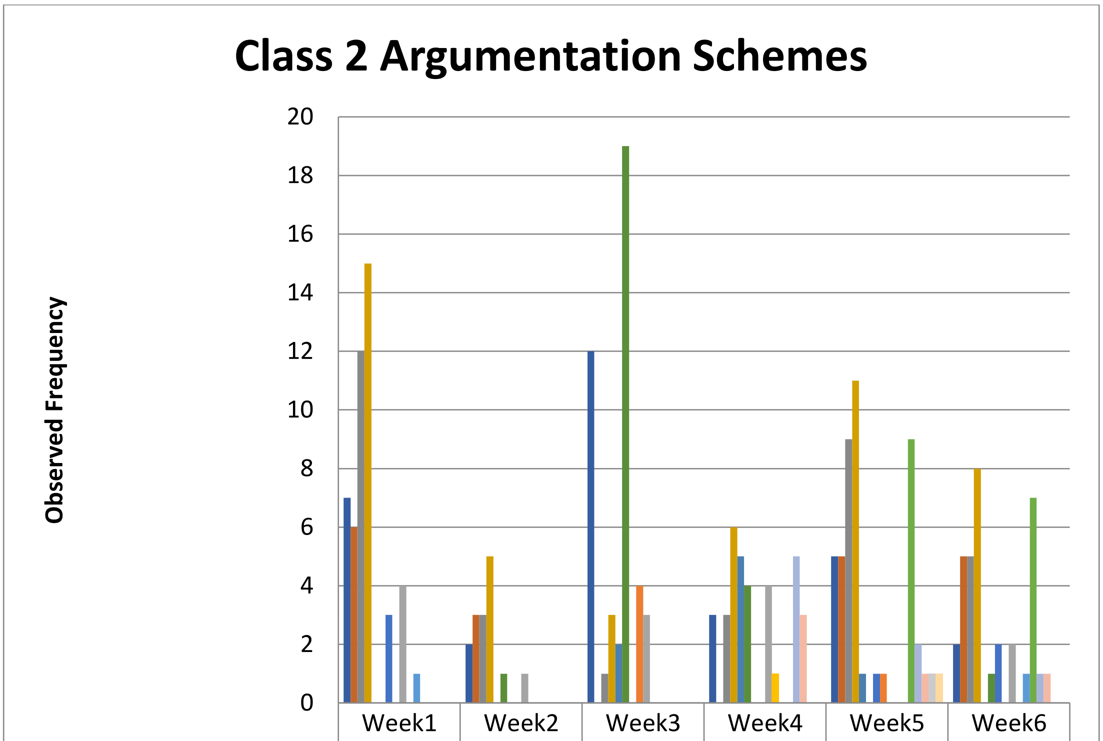
In the last week, students used simulation activity and content of the topic was the factors affecting bulb brightness. In this week, students used mostly argument from evidence to hypothesis (29 %), argument from position to know (29 %), and argument from comparison (18 %). Argument from popularity, cause-effect and analogy were the other argumentation schemes used in this week. Accordingly, the reason why students mainly used argument from evidence to hypothesis and position to know can be linked to nature of argument based inquiry. Accordingly, students conducted their experiments (in simulations) and presented them in negotiation phases. While presenting their results, students mainly pointed out their hypothesis and testing procedure. Hence, they used argument from evidence to hypothesis. While proponent students used argument from evidence to hypothesis when presenting results, other students asked questions and examined the validity of findings. By this way, they actively used argument from position to know. Likewise, students compared different properties of wire such as thickness, length and type. Hence, argument from comparison was the third most used argumentation scheme in last week. Argument from popularity was also used in this week and students used this scheme when they reached the same results with presenting groups. When one argument was proposed in class, others supported it because their argument was same. By this way, proposed argument became popular in class and other students appreciated it. In this part, argumentation schemes occurred in class 1 was presented and next part informs about the argumentation schemes revealed in class 2.

In conclusion, findings about argumentation schemes in Class 1 showed that 13 different argumentation schemes were used in class 1. Total number of argumentation schemes was most in week 3 and week 4 followed this week. On the other hand, students used the fewest number of argumentation schemes in week 6 and week 2 in Class 1. There was no clear pattern to claim that total number of argumentation schemes class 1 students used increased or decreased over time. Although total number of argumentation schemes did not increase over time, findings suggested that total number of types of argumentation schemes that Class 1 students used increased over time in general. The most used argumentation scheme used by Class 1 students were argument from expert opinion. Argument from sign and argument from position to know followed the argument from expert opinion. Argument from comparison, argument from evidence to hypothesis, and argument from corelation to cause were other argumentation schemes frequently used by Class 1 students. In week 1, Class 1 students mostly used argument from position to know, but they mostly used argument from comparison in week 2. On the other hand, argument from expert opinion was the most used scheme in week 3. Likewise, students mostly preferred to use argument from sign in week 4 and week 5. Argument from evidence to hypothesis and argument from position to know were the two argumentation schemes that Class 1 students mostly used in last week of the study.

#### 4.4.2. Class 2 Findings.

Similar to class 1, class 2 students' argumentation schemes were revealed by use of Walton's (1996) argumentation schemes, and Duschl's (2007) ideas about argumentation schemes. Figure 4.7 presents total frequencies of argumentation schemes used by Class 2 throughout the study. Although there are 25 possible argumentation schemes, Class 2 students proposed 16 different argumentation schemes (e.g. argument from evidence to hypothesis) throughout the study and 9 of the argumentation schemes (e.g. argument from gradualism) were not observed in Class 2.

According to Figure 4.7, total number of argumentation schemes was the highest in week 1. Class 2 students used 48 schemes in total in week 1. Students used similar number of argumentation schemes in week 5 and week 3. While they used argumentation schemes 47 times in week 5, they used argumentation schemes 44 times in week 3. Although use of argumentation schemes was not as often as these weeks, students also used them frequently in week 6 and week 4. Total number of argumentation schemes used in week 6 was 35 and in week 4 was week 34. On the other hand, class 2 students did not use argumentation schemes so much in week 2. As a result, weeks can be ranked from highest to lowest in terms of the total number of argumentation schemes occurred as week 1, week 5, week 3, week 6, week 4, and week 2 in class 2.




| | Week1 | Week2 | Week3 | Week4 | Week5 | Week6 |
|---|---|---|---|---|---|---|
| Sign | 7 | 2 | 12 | 3 | 5 | 2 |
| Evidence to Hypothesis | 6 | 3 | 0 | 0 | 5 | 5 |
| Comparison | 12 | 3 | 1 | 3 | 9 | 5 |
| Position to Know | 15 | 5 | 3 | 6 | 11 | 8 |
| Example | 0 | 0 | 2 | 5 | 1 | 0 |
| Expert | 0 | 1 | 19 | 4 | 0 | 1 |
| Popularity | 3 | 0 | 0 | 0 | 1 | 2 |
| Corelation to cause | 0 | 0 | 4 | 0 | 1 | 0 |
| Consequence | 4 | 1 | 3 | 4 | 0 | 2 |
| Committment | 0 | 0 | 0 | 1 | 0 | 0 |
| Cause effect | 1 | 0 | 0 | 0 | 0 | 1 |
| Analogy | 0 | 0 | 0 | 0 | 9 | 7 |
| Established rule | 0 | 0 | 0 | 5 | 2 | 1 |
| Precedent | 0 | 0 | 0 | 3 | 1 | 1 |
| Vagueness | 0 | 0 | 0 | 0 | 1 | 0 |
| Verbal Classification | 0 | 0 | 0 | 0 | 1 | 0 |

*Figure 4. 7* Total frequencies of argumentation schemes used by class 2

Total number of argumentation schemes used in class 2 for each week can also be seen from Figure 4.8 linear line graph:

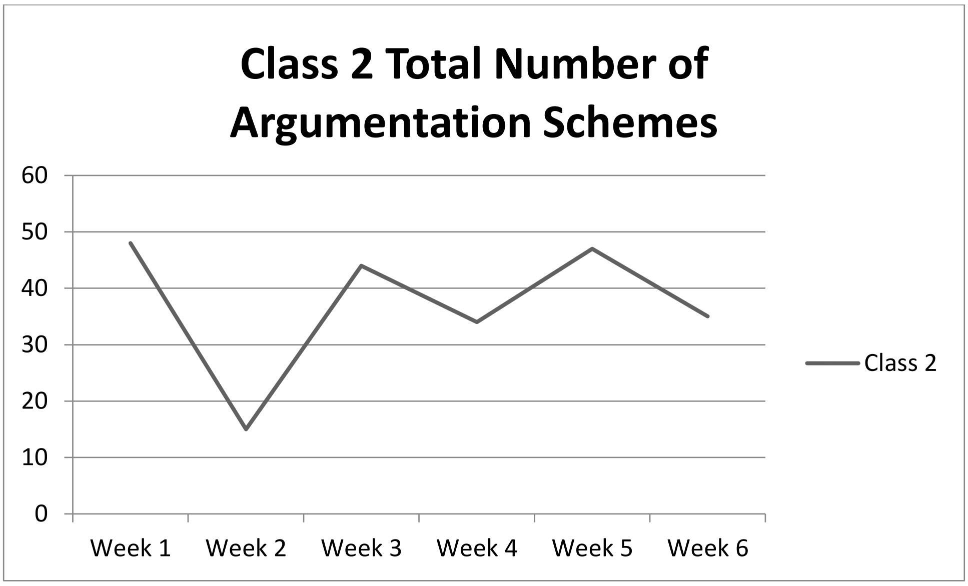



*Figure 4. 8* Class 2 argumentation schemes weekly line graph

Figure 4.7 also provides information regarding the types of argumentation schemes used in each week. Accordingly, total numbers of types of argumentation schemes used in class 2 were the highest in week 5, and students used 12 different schemes which are sign, evidence to hypothesis, comparison, position to know, example, popularity, correlation to cause, analogy, established rule, precedent, vagueness, and verbal classification. Similarly, class 2 students used 11 different schemes in week 6; these schemes used in week 6 are sign, evidence to hypothesis, comparison, position to know, expert, popularity, consequence, cause effect, analogy, established rule and precedent. The week that students used different argumentation schemes third most was week 4. In this week, class 2 students used 9 different argumentation schemes namely sign, comparison, position to know, example, expert, consequence, commitment, established rule and precedent. On the other hand, class 2 students used relatively fewer numbers of different types of argumentation schemes in first three weeks. Accordingly, students used 6 or 7 different types of argumentation schemes in these weeks. In week 1, class 2 students used sign, evidence to hypothesis, comparison, position to know, popularity, consequence, and cause effect. In week 2, similar schemes were used. The schemes used in week 2 by class 2 students were sign, evidence to hypothesis, comparison, position to know, expert, and consequence.

Lastly, class 2 students used sign, comparison, position to know, example, expert, correlation to cause, and consequence in week 3. In conclusion, class 2 students used 6-7 types of argumentation schemes in first three weeks of the study, and they increased the number of types of argumentation schemes in following weeks. The number of types of argumentation schemes ranged 9 and 12 in last three weeks of the study.

Comparison of argumentation schemes according to their use of frequency in class 2 shows that argument from position to know was the most frequently used argumentation schemes in class 2 and students used this scheme 48 times (see Figure 4.7). Second scheme that class 2 students mostly used was argument from comparison which was used 33 times. Beside, argument from sign was the third most frequently used scheme and it was used 30 times. Class 2 students also actively used argument from expert opinion, this scheme was used 25 times in class 2. Argument from evidence to hypothesis, argument from analogy, and argument from consequence were other three schemes that were often preferred by class 2 students. These schemes were used 19, 16 and 14 times respectively. On the other hand, Class 2 students used some argumentation schemes less frequently. For example, argument from example, argument from established rule, argument from popularity, argument from correlation to cause, and argument from precedent were used 8, 8, 6, 5, and 5 times respectively. Examples of other argumentation schemes (e.g., argument from commitment, cause-effect, vagueness, verbal classification) were also available in class 2 whole class discussions, but these schemes were used only once or twice in the study.

Pie charts for each week was also prepared to compare percentages of each scheme occurred in related week. Pie charts including percentages of argumentation schemes occurred in each week for class 2 are presented below as Figure 4.9:

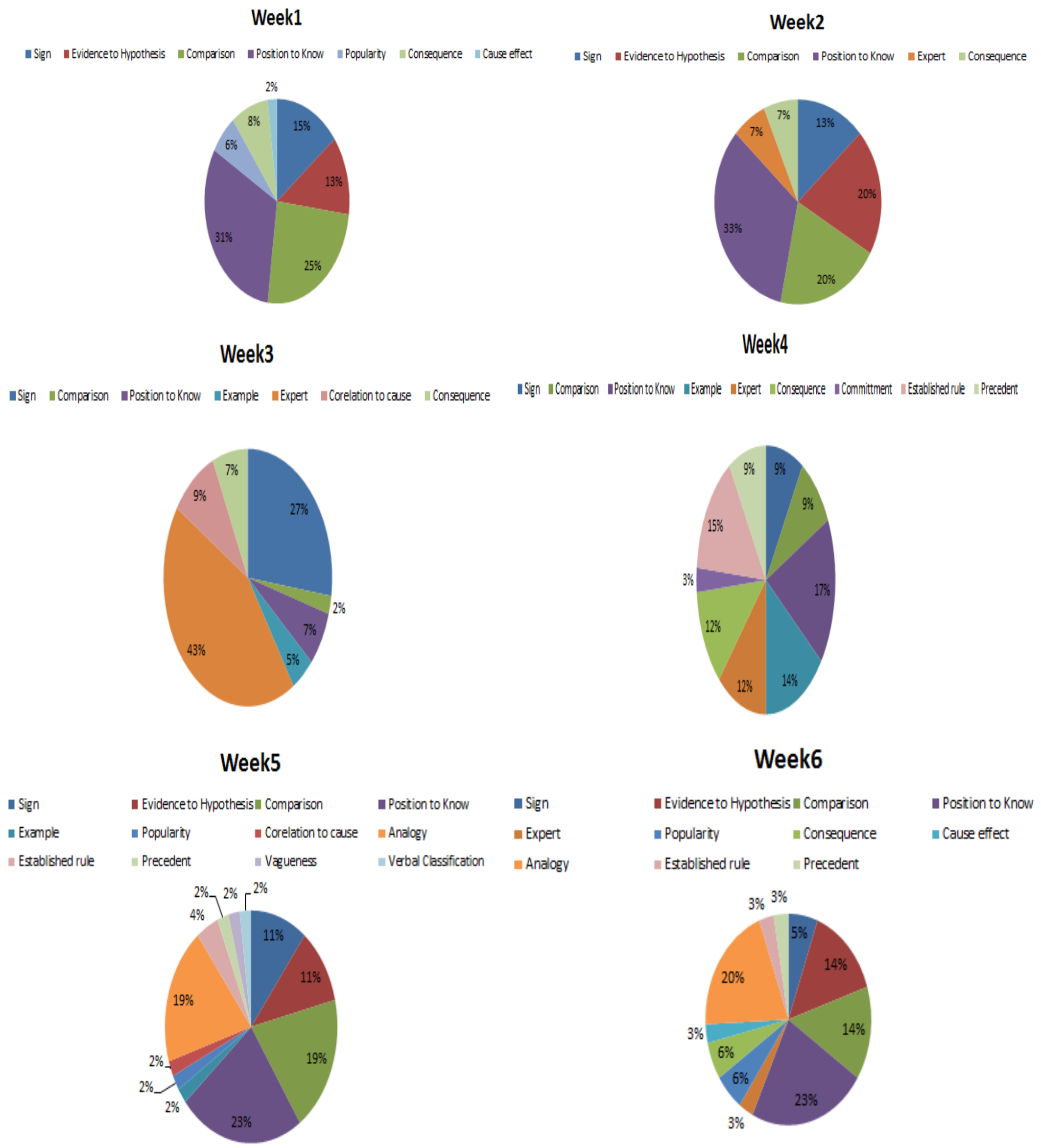



*Figure 4. 9* Class 2 weekly percentages of argumentation schemes

According to Figure 4.9, the most used argumentation scheme was argument from position to know in week 1, the percentage of this scheme was 31 %. This scheme was used when students looked for further information about other groups' research. One quarter of the argumentation schemes were argument from comparison in this week. Accordingly, students preferred to use this scheme when they compared concrete materials such as different kinds of cups. Likewise, 15 % of the schemes were argument from sign. Students formulated ideas for the rationale explanations of their observations and used their inferences when proposing argument from sign. For example, one group found plastic spoon conducted heat better than metal spoon and one student said that “maybe type of the plastic affects the rate of melting”. The fourth most used scheme in week 1 was the argument from evidence to hypothesis, this scheme’s percentage was 13 % and this scheme was used when students presented their experiments and reported their hypothesis and test results. Beside, students used argument from popularity, argument from consequence, and argument from cause-effect in this week, but these schemes percentages were low ranging from 2 % to 8 %.

Similar to first week, class 2 students mostly used argument from position to know in week 2. One third of the total argumentation schemes were argument from position to know in this week. While students discussed with each other, they sought further information and so preferred to use argument from position to know. Likewise, argument from comparison and argument from evidence to hypothesis were the second most used argumentation schemes in week 2. This result was consistent with week 1’ results. The percentages of each scheme were 20 % in week 2. Accordingly, students compared the heat differences of different cups to detect which cup is better heat insulator and so they used argument from comparison. Likewise, students used argument from evidence to hypothesis when they presented their research questions, hypothesis and test results in week 2. Argument from sign was fourth most used argumentation scheme in this week. Similar to week 1, this scheme was used when students tried to make rationale explanations to surprising events such as faster heat transfer of plastic cup. Apart from these schemes, class 2 students also used argument from expert opinion and argument from consequence, but these schemes percentages were low (7 % for each).

Third week’s topic was thermal insulating product and students used evidence cards to construct their arguments. Moreover, it should be added that students had no prior knowledge about this topic. Students consulted on the information found in evidence cards and so percentages of argument from expert opinion reached to 43 % in this week. The second most used argumentation scheme in week 3 was argument from sign. Although students had no prior knowledge to match related thermal insulating products and parts of house, they still made inferences using their daily life experiences and tried to match insulating products and relevant house parts. For example, when one student learnt that silicone wool decrease the water leakage, she claimed that silicone wool might be used for windows because windows are affected by rain. Such inferences derived from daily life experiences might cause students frequently use argument from sign scheme in this week. Students also used other schemes in this week; however, percentages of these schemes were highly low comparing with the two schemes (i.e., argument from expert opinion and argument from sign). Accordingly, 9 % of argumentation schemes were argument from correlation to cause. This scheme was used when students used wrong cause effect relationships. For example, one student claimed that wood wool can be used for exterior wall as insulating material because wood wool is environmentally friendly. There was no plausible link between using wood wool for exterior wall and the concept “environmentally friendly” so this argument was used as argument from corelation to cause. Students also seldom considered the possible results of the actions (realization of claims), and so used argument from consequence (7 %). Similarly, students asked further information from the other groups and used argument from position to know, but percentage of this scheme was low (7 %). Other schemes that students used in this week were argument from example and argument from comparison; however, their percentages were highly low comparing with schemes explained above.

In week 4, topic of whole class discussion was energy sources. Students both had prior knowledge about topic and evidence cards were provided them. By using their prior knowledge and information found in evidence cards class 2 students constructed and examined many arguments in this week. Likewise, class 2 students used numerous types of argumentation schemes. Therefore, it is possible that percentages of argumentation schemes are very close to each other. In other word, none of the argumentation schemes dominated this week. According to Figure 4.9; outstanding argumentation schemes for week 4 were argument from position to know (17 %), argument from established rule (15 %), argument from example (14 %), argument from expert (12 %) and argument from consequence (12 %).

Moreover, students used argument from sign, argument from comparison, and argument from precedent same amount. Each of these argumentation schemes corresponded to 9 % of all schemes occurred in week 4. Students used argument from position to clarify presenting group's reasoning about their selection of energy sources. For example, one of the groups chose not to use wood as energy source because wood is non-renewable. After this explanation, one of the students asked this group why they did not report other fossil fuels as the energy sources they did not prefer because other fossil fuels are also non-renewable energy sources. Similar to argument from position to know, students used argument from established rule. Established rule argument examples are about the rules that human beings constructed in Walton's (1996) explanations. However, class 2 students used natural laws in their arguments and these arguments matched with argument from established rule. For example, one of the students selected solar energy as energy sources. While explaining the advantages of this energy, he said that we can always find solar energy because Earth rotates and always one side of the Earth gets solar energy. As seen from this example, students can use natural laws as their source of reasoning in their arguments. Students also used argument from example. This scheme is used when people aim to support their generalization. For example, students preferred to use renewable energy source which was their general idea. Then, they gave examples about types of different renewable energy sources such as geothermal energy and wind energy and their benefits to support their general idea. Likewise, students used evidence cards to construct their arguments. By referring evidence cards, they used argument from expert opinion. Moreover, students considered possible costs and profits of each energy types when they selected them. While students determine possible costs and profits, they actively used argument from consequences in this week.

Students started electricity topic in week 5 and they conducted experiments regarding electrical conductivity. Students used 12 different schemes in this week. Five of these schemes were used more often than other seven argumentation schemes. The more frequently used argumentation schemes in week 5 are argument from position to know (23 %), argument from analogy (19 %), argument from comparison (19 %), argument from sign (11 %) and argument from evidence to hypothesis (11 %). On the other hand, less frequently used argumentation schemes used in this week are argument from example, argument from popularity, argument from correlation to cause, argument from established rule, argument from precedent, argument from verbal classification and argument from vagueness of verbal classification. The percentages of these seven schemes were not noticeable and ranged between 2 % and 4 %. Students used argument from position to know by asking questions to presenting groups to clarify content. They also asked questions to understand the reasons of their observations. For example, students used argument from position to know to understand why salty water color turned to yellow when circuit cables were touched to salty water. Students also actively used argument from comparison. In their experiments, students compared different materials in terms of their electrical conductivity. For example, they compared conductivity of salty water, sugared water and tap water in their experiments. While reporting their results, they often used argument from comparison. Students not only compared concrete objects, they also compared abstract themes, and so they actively used argument from analogy in their experiments. For example, they linked between heat conductivity and electrical conductivity, so they frequently used argument from analogy. Unfortunately, the content of argument from analogy was wrong because there is no link between heat conductivity and electrical conductivity according to their scientific explanations. In this week, students also actively used argument from sign. This scheme was mostly used when students tried to explain why salty water conducts electricity. They used their daily life observations in their inferences and most of their explanations were not consistent with scientific explanation. However, they were aware that their ideas can be wrong, all of their sentences included possibility. They were just making predictions when using argument from sign. For example, one student thought that maybe copper in wire leaked into salty water, and copper changed the water color into yellow. This explanation was wrong, however, student was aware that this explanation is a possibility because student started the sentence with “maybe”. Likewise, students used argument from evidence to hypothesis in this week. This argument was mainly used when students reported their hypothesis, research questions and results.

Last week topic was the factors affecting bulb brightness in electrical circuit. Students used simulations to conduct their experiments in this week. Totally, 11 different argumentation schemes were used in this week. Four of the schemes were more often used than seven argumentation schemes. The four more frequently used argumentation schemes in this week are argument from position to know (23 %), argument from analogy (20 %), argument from evidence to hypothesis (14 %) and argument from comparison (14 %). The less frequently used seven argumentation schemes in this week, on the other hand, are argument from sign, argument from expert, argument from popularity, argument from consequence, argument from cause effect, argument from established rule, and argument from precedent. The percentages of these less frequently used argumentation schemes ranged between 2 % and 6 %. As in all weeks except week 3, argument from position to know was most used argumentation scheme in week 6. When using argument from position to know, students asked questions about the results and procedures of experiments. For example, one group examined the wire thickness effect on bulb brightness. After group presented their results, one of the students from other groups asked whether this group thickened the plastics layer that wrapped wire or just thickened the metal wire in their experiment. Similar to week 5, class 2 students frequently used argument from analogy. Some of the analogies that students used in these arguments were correct and some others were false. Arguments from analogies which were correct were about the core idea of this week. Accordingly, some students linked the length of wire with the long distance. In such analogies, students claimed that it is difficult to pass long distance comparing with short distance. Likewise, long wire resists more to current comparing with short wire. These kinds of analogies were also used by teachers and they are correct. On the other hand, students still used some false analogies. Accordingly, they still linked heat conductivity and electrical conductivity in their analogies when they tried to explain reason of electrical conductivity. Students also frequently used argument from comparison and argument from evidence to hypothesis in this week. Students compared different types of wires, lengths of wires and thickness of wire in their experiments. By doing this, they often used argument from comparison. Likewise, students prepared their hypothesis and tested them. When they presented their hypothesis and results, they used argument from evidence to hypothesis. In conclusion, this part separately presented findings of class 1 and class 2 regarding argumentation schemes students used. The next part shows the assertions about argumentation schemes. These assertions are derived from either similarities or differences between class 1 and class 2. It is thought that these assertions contribute on enlightening of nature of argumentation schemes that middle school students use.

In conclusion, findings about Class 2 students' argumentation schemes show that 16 different argumentation schemes were used in class 2. Total number of argumentation schemes was most in week 1, week 5 and week 3. On the other hand, students used the fewest number of argumentation schemes in week 2. There was no clear pattern to claim that total number of argumentation schemes class 2 students used increased or decreased over time. Although total number of argumentation schemes did not increase over time, findings suggested that total number of types of argumentation schemes that class 2 students used increased over time in general. The most used argumentation scheme used by class 2 students was argument from position to know. Argument from comparison and argument from sign followed argument from position to know in class 2. Other frequently used argumentation schemes in class 2 were argument from expert opinion, argument from evidence to hypothesis, argument from analogy, and argument from consequence. Weekly comparison of argumentation schemes via pie charts further informed that most preferred argumentation schemes in Class 2 was argument from position to know in all weeks except week 3. In week 3, the most preferred argumentation scheme was argument from expert opinion in class 2.

### 4.4.3. Assertions about argumentation schemes derived from similarities between class 1 and class 2.

Class 1 and class 2 had some similarities regarding their use of argumentation schemes throughout the study. Based on these similarities, three assertions have been proposed. These assertions are explained below:

Assertion 1: The number of types of argumentation schemes used by the students increases over time.

This study showed that both class 1 and class 2 increased the total number of type of argumentation schemes over time. These findings may suggest that students do more reasoning over time and they use wider range of argumentation schemes.

More specifically, class 1 used five different argumentation schemes at the beginning of the study and the number of type of argumentation schemes reached eight in week 5. Class 1 used argument from position to know, argument from comparison, argument from evidence to hypothesis, argument from sign, and argument from example in week 1. However, in week 5, class 1 students added three more argumentation schemes to these five argumentation schemes. The additional schemes used by class 1 students in week 5 are argument from correlation to cause, argument from cause effect, and argument from gradualism. For example, argument from gradualism was used in week 5, but this scheme was not observed in previous weeks. According to Walton (1996), people are not persuaded by use of one comprehensive argument. Therefore, more than one argument which are linked but less comprehensive are used to persuade others in argument from gradualism. One of the students from class 1 explained the relationship between distance of cable ends in salty water and brightness of bulb step by step using argument from gradualism. The following example shows this argument from gradualism:

Student 132: Cable that is connected to battery gets current from battery and electricity reach in salty water from the end of this cable. Electricity should pass to other cable's end to light the bulb. When we decrease the distance between ends of cables, more current will pass and less energy will lose in salty water. The more current passes, the brighter bulb we get.

If this student had just said that there is a reverse relationship between distance of cables in salty water and bulb brightness because of the resistance or electricity flow, other students would not have been persuaded. On the other hand, step by step explanation as argument from gradualism facilitated other students' understanding the relationship between distance of cables in salty water and bulb brightness and these students most probably accepted this explanation.

Likewise, class 2 started study with seven different argumentation schemes, and this class used 12 different argumentation schemes in week 5. According to this, class 2 students used argument from position to know, argument from comparison, argument from sign, argument from evidence to hypothesis, argument from popularity, argument from consequence, and argument from cause effect in week 1. On the other hand, this class added some new argumentation schemes to the existing argumentation schemes in week 5. These argumentation schemes added in week 5 are argument from analogy, argument from example, argument from correlation to cause, argument from established rule, argument from precedent, argument from verbal classification and argument from vagueness. For example, class 2 students did not use argument from analogy until week 5. However, argument from analogy became the second most used argumentation scheme in week 5. Its percentage was 19 %. For example; one of the students used an analogy between pure water and sugared water while explaining why sugared water does not conduct electricity.

Student 231: We know that pure water does not conduct electricity because it is neutral. Likewise, when we use sugared water in circuit, it does not conduct electricity too. The reason why sugared water does not conduct electricity is the same with pure water. Sugared water is neutral like pure water.

Assertion 2: Students prefer to use argument from expert opinion, argument from sign, argument from position to know, argument from comparison, and argument from evidence to hypothesis more often than other argumentation schemes.

Findings of the study showed that students prefer to use some argumentation schemes more than others in middle school level science lessons. Although there are 25 argumentation schemes according to Walton (1996), Both class 1 and class 2 students most frequently used five argumentation schemes which are argument from expert opinion, argument from sign, argument from position to know, argument from comparison, and argument from evidence to hypothesis. Accordingly, class 1 students used these schemes 41, 40, 23, 22, and 19 times respectively. Likewise, class 2 students used these schemes 25, 30, 48, 33 and 19 times.

Students generally used argument from expert opinion when they used second hand data to construct arguments. Both classes used frequently argument from expert opinion when evidence cards are provided them. Examples of argument from expert opinion are presented below:

Class 1

Student 124: We should not use tar as insulating product in interior wall because it is flammable, not durable and used for ceiling (Student is reading evidence card)

Class 2

Student 241: I think we can use wood wool in interior wall as insulating product because evidence card provides this information.

Likewise, argument from sign was frequently used by both classes. Argument from sign was used for different aims throughout the study. For example, students used this scheme when they looked for explanation of the observed phenomena. In the following example; one student from class 1 tried to explain why salty water conducts electricity. In this example, student did not know the ion and their effect on electrical conductivity. Therefore, argument from sign he used mislead him when he tried to explain reason of observation:

Student 115: the reason why salty water conducted electricity in your experiment, can be related with your experimental design. You might connect the cable ends each other in salty water and therefore, bulb might light in your experiment.

Third scheme that was frequently used by both classes is argument from position to know. Argument from position to know was mainly used by two ways. Students either used this argumentation schema to get further information or they used this scheme to show inconsistencies that proponents of the idea had. The following first excerpt is example of using argument from position to know to get more information and second excerpt is an example of using argument from position to know to show inconsistencies to the proponent.

Student 114: In which cup, did you conduct your experiment? (Argument from position to know-to obtain more information)

Student 224: We heated spoons in hot water. After that we removed spoons from water and measured their temperature using thermometer.

Student 213: How did you measure solid substance's temperature using thermometer? Thermometer is used to measure liquids or gases' temperatures. (Argument from position to know-to refer inconsistencies).

Similarly, both classes used argument from comparison throughout the study. This scheme was used when students compare concrete materials with each other. When students explain their research design, they frequently compared different materials in terms of heat conductivity or electrical conductivity. By comparing different materials in their experiments, students frequently used argument from comparison. The followings are some of the examples showing that students used argument from comparison when they presented their experiments.

Class 1

Student 132: We put different spoons in hot water which was 70˚ about two minutes. We placed butter on each spoon and butter on metal spoon melted first. Therefore, metal conduct heat better than plastic and wood spoon.

Class 2

Student 231: We put spoons on hot water. In our experiment, firstly butter on plastic spoon melted. Secondly, butter on metal spoon melted and lastly, butter on wood spoon melted. Our observations show that plastic spoon conduct heat better.

Finally, both classes of students used commonly used argument from evidence to hypothesis. This argumentation scheme that depends on testing was usually used by students when they reported their experiments and hypothesis. This situation is expected in argument based inquiry study because groups are expected to present their findings with others in negotiation phase of argument based inquiry research. Some examples that students actively used argument from evidence to hypothesis are given below:

Class 1:

Student 131: Our research question was whether type of wire affects the wire resistance. We hypothesized that type of wire affects wire resistance.

Class 2:

Student 243: Our research question was whether thickness of wire affects wire resistance. We thought that it affects. Our control variable is number of bulb. Independent variable is thickness of wire and dependent variable is bulb brightness.

Assertion 3: Type of data students used may affect which argumentation schemes they use in argumentation.

Similarities between two classes showed that type of data students used might affect their selection of argumentation schemes. Accordingly, students used first hand data in first two weeks and last two weeks because they conducted experiment and produced their own data in these weeks. On the other hand, students used second hand data in third and fourth weeks when they obtained information from evidence cards.

When students used first hand data, both classes frequently used argument from evidence to hypothesis. This finding is reasonable, because students needed to test their hypothesis when they conducted experiments. After students conducted experiments, they presented experiment results. In presenting experiment results, they actively used argument from evidence to hypothesis because they reported whether their hypothesis was rejected or retained. However, students did not use argument from evidence to hypothesis when information was directly provided to them as second hand data (i.e., evidence cards) because students did not test their hypothesis or conduct experiments in these weeks.

On the other hand, students most frequently used argument from expert opinions when they used second hand data. Although students actively used argument from expert opinion when they used second hand data, this scheme was not used by students when first hand data (i.e., experiment results) was used. In conclusion, it can be claimed when first hand data replaced with second hand data in argumentation, students might give up using argument from evidence to hypothesis and start to use argument from expert opinion. Likewise, when students stop using second hand data and start using first hand data, they may give up using argument from expert opinion, and start to use argument from evidence to hypothesis. In this part, assertions about argumentation schemes derived from similarities between class 1 and class 2 are presented. The next part presents the assertions about argumentation schemes derived from differences between class 1 and class 2.

#### 4.4.4. Assertions about argumentation schemes derived from differences between class 1 and class 2.

As classes had some similarities about argumentation schemes, they also had some differences. Based on these differences, two assertions are proposed about argumentation schemes. These assertions are presented in this part.

Assertion 1: More active class uses more argumentation schemes comparing with less active class.

Findings of the study showed that class 2 engaged in argumentation process more than class 1. Accordingly, class 2 was labelled as more active class and class 1 was labelled as less active class. As more active class engaged in argumentation process more than less active class, more active class used more argumentation schemes comparing with less active class throughout the study. Actually, two classes did not separate from each other first three weeks. Classes used 5-7 different argumentation schemes in first three weeks. However, number of use of different argumentation schemes differed after week 4. While less active class used up to eight different argumentation schemes after week 4, more active class used up to twelve argumentation schemes in the following weeks. Total numbers of types of argumentation schemes used by two classes throughout the study are presented in Figure 4.10;

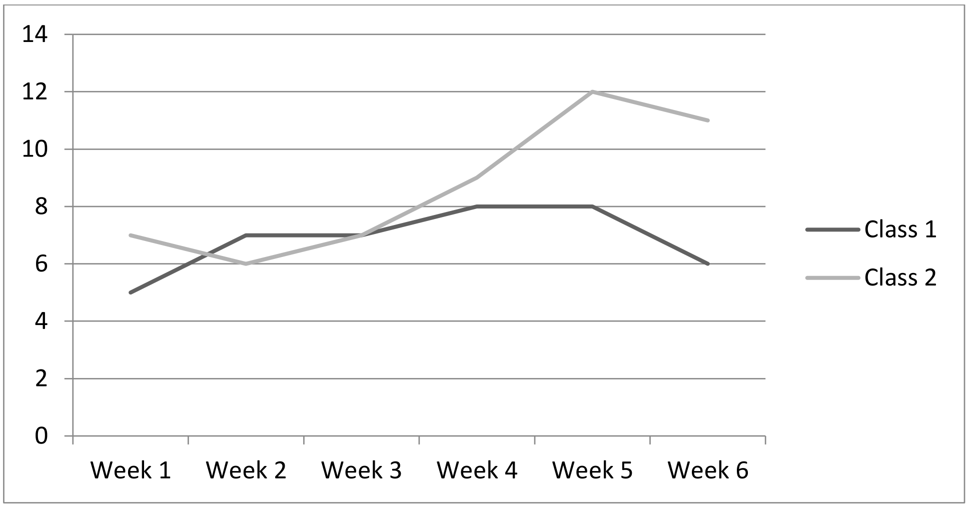



*Figure 4. 10* Total numbers of types of argumentation schemes used by two classes

These findings supported the idea that when students engaged in argumentation more, they make more reasoning and they appeal to various types of argumentation schemes.

Assertion 2: While the most used argumentation schemes is argument from position to know in more active class, the most used argumentation schemes is argument from expert opinion in less active class.

Classes also differed regarding the most used argumentation schemes. The most used argumentation scheme in less active class was argument from expert opinion which was used 41 times; on the other hand, this scheme was used 25 times in more active class. While most used schema was argument from expert opinion in less active class, more active class mostly used argument from position to know which was used 48 times. On the other hand, this schema was used 23 times in less active class.

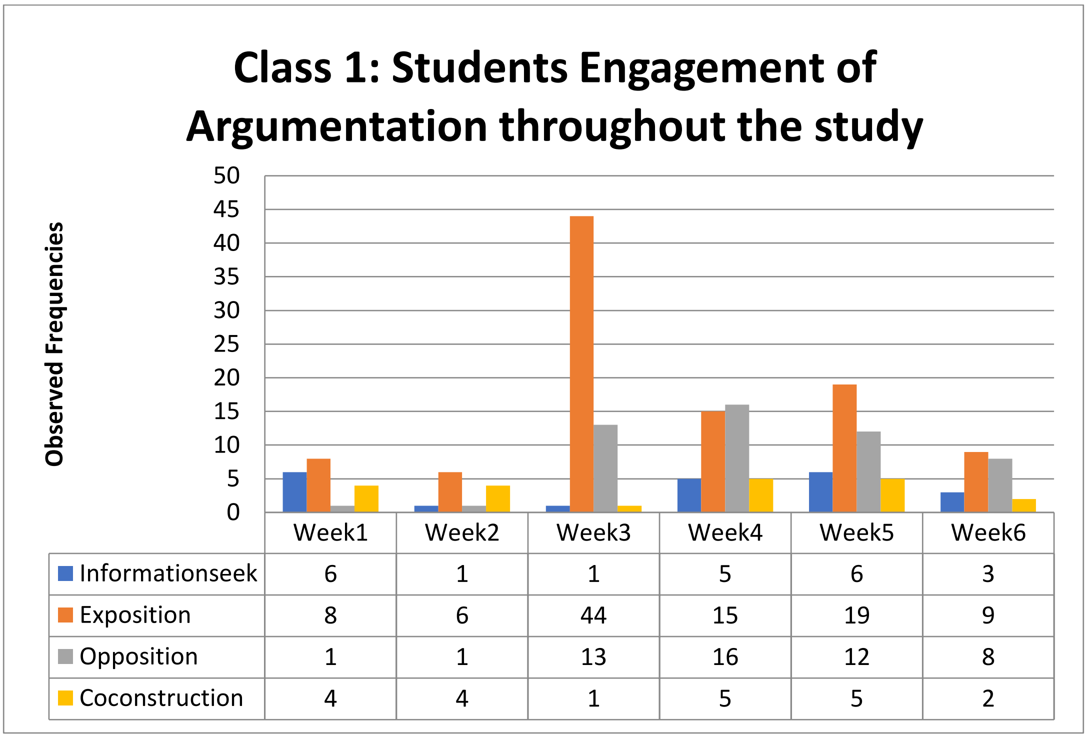
In line with this, it can be claimed that less active class (i.e., Class 1) appeal to second hand data when producing argument, so they use evidence found in second hand data and use argument from expert opinion. On the other hand, more active class (i.e. Class 2) focused on asking more questions to others during argumentation and they mainly preferred to use argument from position to know. Examination of others' arguments and experiments through argument from position to know might facilitate to sustain argumentation in Class 2 and make this class as more active class comparing with the other class (i.e. Class 1) which mainly relied on expert opinion when constructing arguments. The next part presents findings about students' engagement in argumentation process.

## 4.5. Findings for the Nature of the Students' Engagement in Argumentation Process

In order to address the fifth research question of the study, namely "What is the nature of 6$^{th}$ grade students' engagement in argumentation process in ABI treatment?", Sampson and Clark's (2011) codes (e.g., exposition) were used. In line with this, total frequency of codes (i.e., engagement components), linear line graphs and their percentage were reported for both classes in each week. Firstly, findings of class 1 were reported in this part and class 2 findings follow it. After findings of two classes were reported, assertions derived from the similarities of both classes' findings, and assertion derived from differences of both classes' findings was presented to address the research question about nature of students' engagement in argumentation process.

### 4.5.1. Class 1 Findings.

The study lasted six weeks and whole class discussions were deductively analysed for class 1. As detailed in the methodology chapter (chapter 3), the deductive codes used to reveal students' engagement in argumentation processes were Information seeking, Expositional Comment, Oppositional Comment and Co-construction of knowledge. These codes were obtained from Sampson and Clark's (2011) study. Based on this analysis, total frequency of four engagement components was obtained per each week. Figure 4.11 shows the total frequency of four engagement components revealed in class 1's whole class discussions. Frequency of these engagement components occurred in every week represents students' engagement in argumentation process.




| | Week1 | Week2 | Week3 | Week4 | Week5 | Week6 |
|---|---|---|---|---|---|---|
| Informationseek | 6 | 1 | 1 | 5 | 6 | 3 |
| Exposition | 8 | 6 | 44 | 15 | 19 | 9 |
| Opposition | 1 | 1 | 13 | 16 | 12 | 8 |
| Coconstruction | 4 | 4 | 1 | 5 | 5 | 2 |

*Figure 4. 11* Summary of class 1 students' engagement of argumentation process

According to Figure 4.11, class 1 students engaged in argumentation process at most in week 3 which is about thermal insulating products. Students engaged in argumentation 59 times in this week. The next highest frequency reflecting students' engagement was observed in week 5 week when the topic was electrical conductivity. Students engaged in argumentation process 42 times in this week which was comparable to that of week 4: students engaged in argumentation 41 times in total in week 4 when the topic was types of fuels. While students were relatively active in week 3, week 4 and week 5, they less actively participated in argumentation in week 1, week 2 and week 6. More specifically, students engaged in argumentation process 22 times in week 6 when the factors affecting the bulb brightness was covered, On the other hand, students participated in argumentation process 19 times in week 1 when the topic was heat conductors and they engaged in argumentation process 12 times in week 2 when the heat insulators were covered. In conclusion, the order of the students' engagement in the argumentation process from the highest to the lowest based on the overall frequencies across weeks was week 3, week 5, week 4, week 6, week 1 and week 2

When the four engagement components were compared based on their overall observed frequencies throughout the study, Expositional comment emerged as the component which was used at most (see Figure 4.11). This engagement component was utilized 101 times. Oppositional comment, which was used 51 times, was the second most used engagement component. Information seeking and Co-construction of knowledge were used relatively less. More specifically, information seeking was used 22 times and co-construction of knowledge was used 21 times.

While frequency graph (figure 4.11) informed about students' total number of engagement in argumentation process, pie charts were used to compare four engagement components with each other for each week. Pie charts show percentages of each engagement component occurred in each week and are presented in figure 4.12.

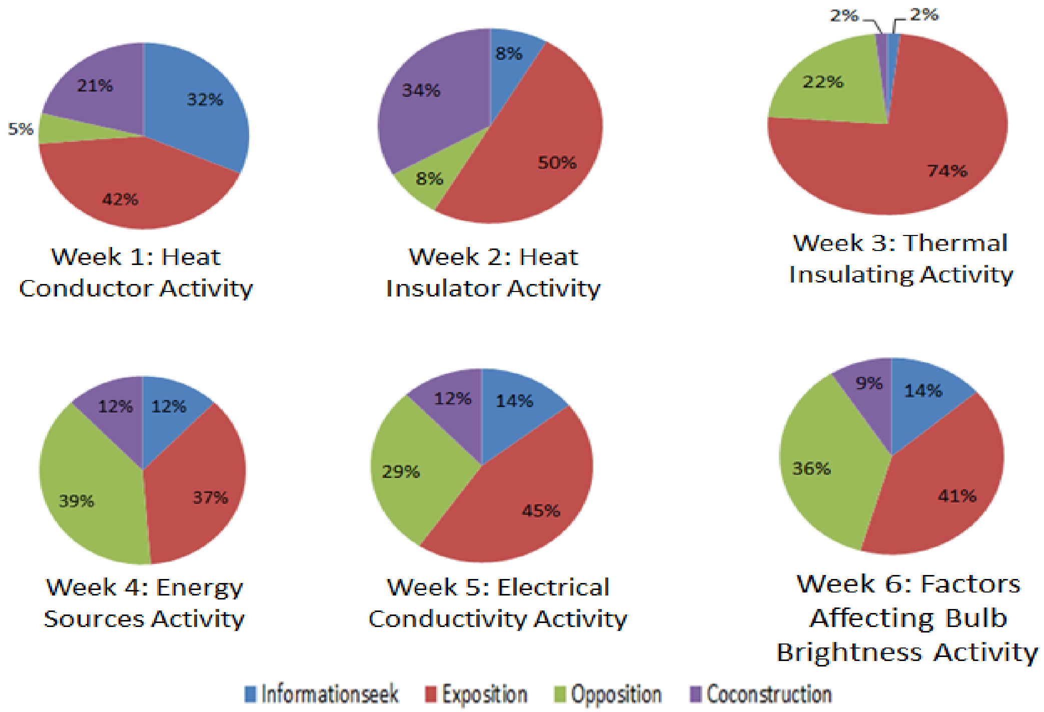



*Figure 4. 12* Pie charts showing the percentage of engagement components for class 1

Figure 4.12 shows that the most used engagement component for Class 1 students' engagement in argumentation process was expositional comment in all weeks except week 4. Accordingly, nearly half of the argumentation was expositional comment in all weeks. Percentages of expositional comment ranged between 37 % and 74 %. On the other hand, the engagement component that was used at most in week 4 was oppositional comment with 39 %.

Figure 4.12 also shows that the least used engagement component was oppositional comment in week 1 and week 2. In week 3 and 4, on the other hand, the least used engagement components were information seeking and co-construction of knowledge while in week 5 and 6, the least observed component was co-construction of knowledge.

Apart from total number of frequencies and pie charts, linear line graphs were used to detect possible changes in individual engagement components representing students' engagement in argumentation across weeks. These line graphs are presented in Figure 4.13;

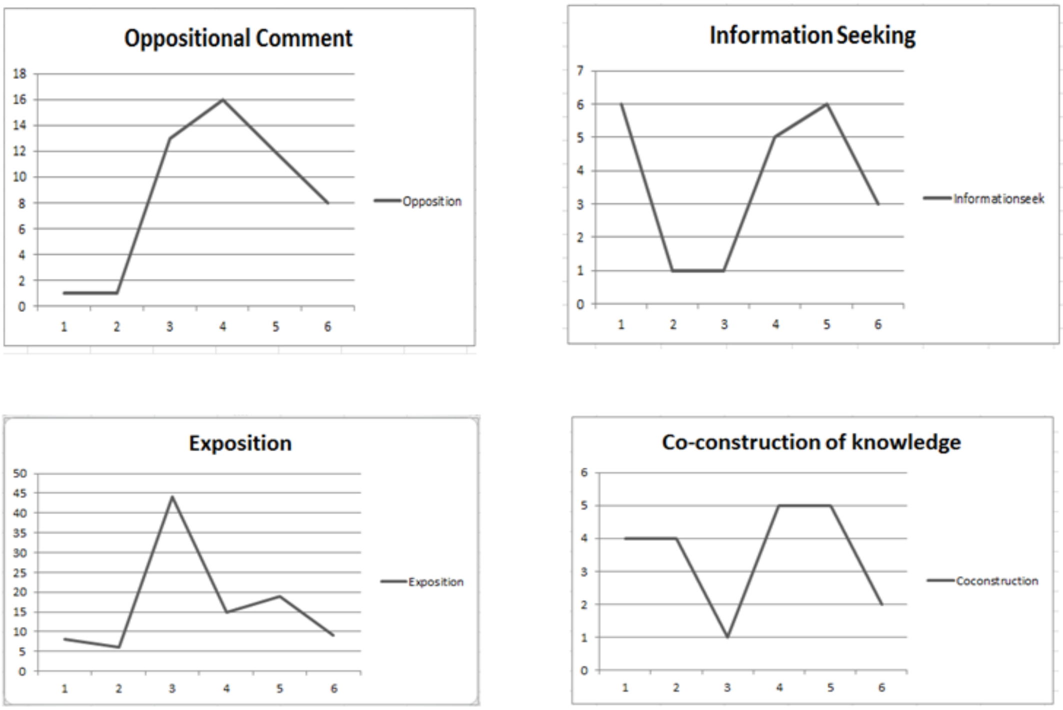



*Figure 4. 13* Weekly changes of engagement components for class1

According to Figure 4.13, oppositional comment was the unique engagement component whose number of frequency increased at the end of the study comparing with first week of the treatment. While it was rarely used in first two weeks, students started to use it more in the following weeks. However, same trend was not observed for other three engagement components. Data shows that there is a fluctuation of the frequency for these three engagement components in the study.

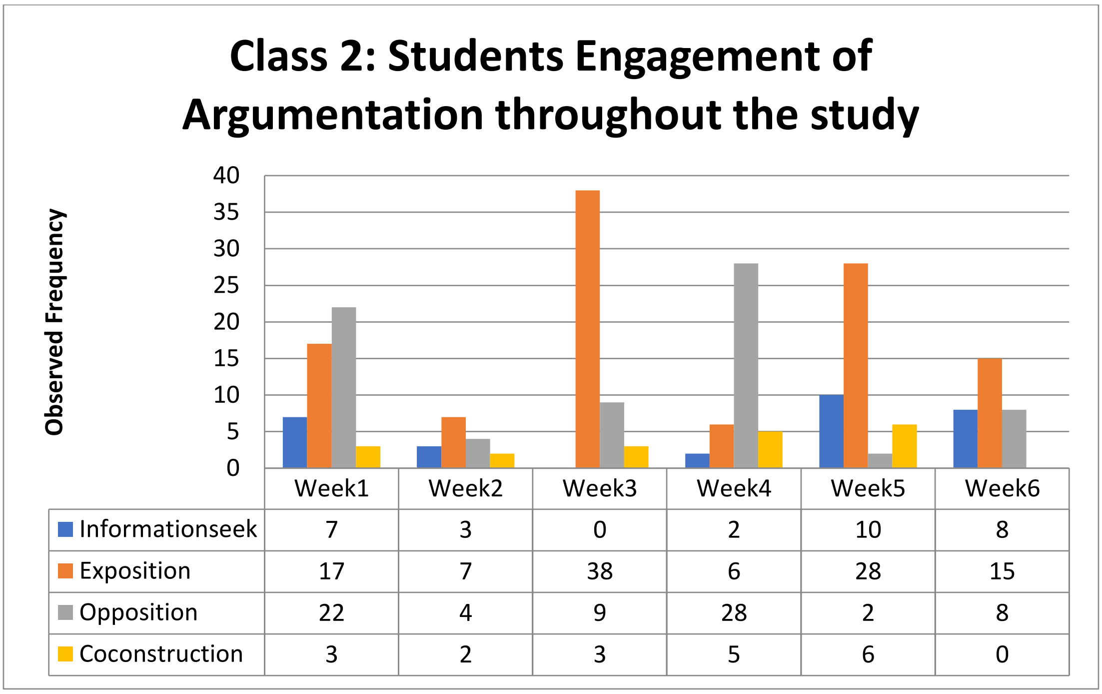
Figure 4.13 shows that students used oppositional comment at most in week 4 that is about types of fuels (16 times). On the other hand, students used oppositional comments least in first two weeks. Students used this engagement component just one time in these weeks. On the other hand, linear line graph for information seeking shows that students used information seeking at most in week 1 and week 5. They used information seeking 6 times in these weeks. There was one common point between week 1 and week 5 that both weeks were the first weeks of their unit. While week 1 was first week of matter and heat unit, week 5 was first week of electricity unit. On the other hand, information seeking was used only one time in week 2 and week 3. Next, line graph for the change of exposition component shows that students used exposition comment at most in week 3 (44 times) and they used it least in week 2 (6 times). Figure 4.13 also shows the changes of co-construction of knowledge throughout the study. According to linear line graph for co-construction of knowledge, class 1 students hardly used co-construction of knowledge in week 3 (1 time) and week 6 (2 times). Students relatively used co-construction of knowledge more in other weeks (4-5 times).

In sum, the findings suggested that the most used engagement component was appeared to be expositional comments in class 1. Expositional comment was followed by oppositional comment. On the other hand, information seeking and co-construction of knowledge was observed less frequently in this class. Although frequency of oppositional comments increased at the end of study comparing with first week, other three components representing students' engagement did not obviously increase or decrease. On the other hand, these three components fluctuated through weeks.

Up to this point, class 1 findings about students' engagement in argumentation process were presented and next part informs about the findings of class 2 students' engagement in argumentation process.

#### 4.5.2. Class 2 Findings.

The same engagement components were used to reveal Class 2 students' engagement in argumentation process obtained from whole class discussions throughout the study. Total frequency of each engagement component was presented weekly in Figure 4.14.




| | Week1 | Week2 | Week3 | Week4 | Week5 | Week6 |
|---|---|---|---|---|---|---|
| Informationseek | 7 | 3 | 0 | 2 | 10 | 8 |
| Exposition | 17 | 7 | 38 | 6 | 28 | 15 |
| Opposition | 22 | 4 | 9 | 28 | 2 | 8 |
| Coconstruction | 3 | 2 | 3 | 5 | 6 | 0 |

*Figure 4. 14* Summary of class 2 students' engagement of argumentation process

According to Figure 4.14, class 2 students engaged in argumentation process at most in week 3 (thermal insulating products). Students engaged in argumentation 50 times in this week. Unlike class 1, total frequency of engagement components per each week was similar in class 2. For example, the next highest frequency reflecting students' engagement was observed in week 1 for class 2. In this week, students engaged in argumentation 49 times. Similarly, class 2 students engaged in argumentation 46 times in week 5 and they engaged in argumentation 41 times in week 4. On the other hand, students were less active in week 6 and week 2. Accordingly, students engaged in argumentation 31 times in week 6. Class 2 students were least active in week 2, and students engaged in argumentation 16 times in this week.

Comparison of Figure 4.11 and Figure 4.14 clearly suggested that class 2 students engaged in argumentation process generally more than class 1 because total number of frequency of engagement components revealed in Class 2 are more than Class 1 for each week except for week 3 and week 4.

Figure 4.14 shows which engagement component was used more than others. Accordingly, Expositional comment was the most used engagement component and it was used 111 times throughout the study. Second most used engagement component was oppositional comment and this component was used by class 2 students 73 times. On the other hand, information seeking was the third component in terms of total number of frequency; this component was used 30 times in class 2. The engagement component that was used in fewest numbers in class 2 was co-construction of knowledge which was used 19 times. Although sorting of engagement components based on total number of frequency was the same for class 1 and class 2 (expositional comment >oppositional comment> information seeking> co-construction of knowledge), the number of times that they were used was more in class 2 than class 1 for all four engagement components. Therefore, it can be said that class 2 students were more active than class 1 students when they engaged in argumentation.

While figure 4.14 informs about frequency of engagement components for each week, pie charts were used to compare four engagement components occurrence with each other for each week using their percentage. Pie charts showing percentage of each component for each week revealed in class 2 are presented in figure 4.15;

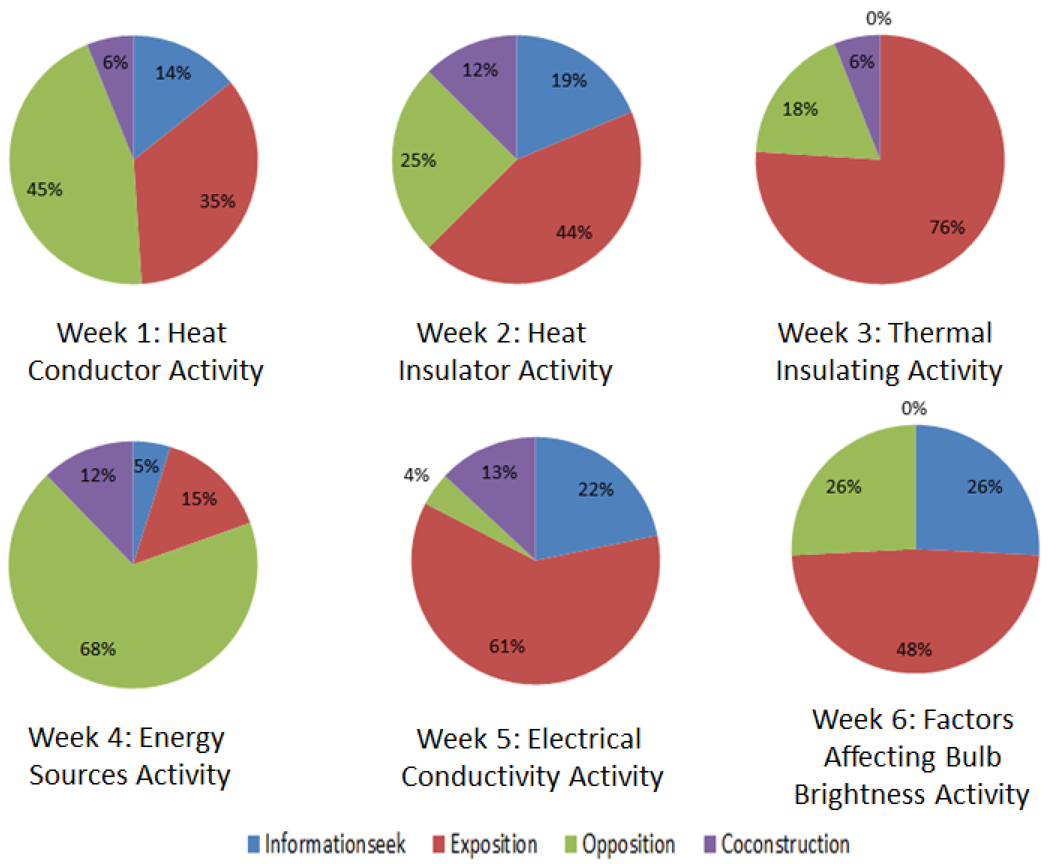



*Figure 4. 15* Pie charts showing the percentage of engagement components for class 2

Figure 4.15 suggests that most used engagement component in class 2 was expositional comments except for week 1 and week 4. Percentage of the expositional comment in the four weeks where it was observed at most ranged from 44 % to 76 %. On the other hand, the engagement component that was used at most in week 1 and week 4 was oppositional comment. Oppositional comment percentage was 45 % in week 1 and 68 % in week 4.

Pie charts also inform the least used engagement components in terms of percentage. Accordingly, the least used engagement component in week 1, week 2 and week 6 was co-construction of knowledge. Its percentages in these weeks were 6 %, 12 % and 0 respectively. On the other hand, information seeking was the least used engagement component in week 3 and week 4. The percentage of information seeking was 5 % in week 4 and it was not used in week 3. On the other hand, oppositional comment was the least used engagement component in week 5. In this week, only 5 % of the total argumentation process was oppositional comments.

Linear line graphs were also used to detect how individual engagement components changed throughout the study. Linear line graphs representing individual change of engagement components happening in class 2 are presented in Figure 4.16;

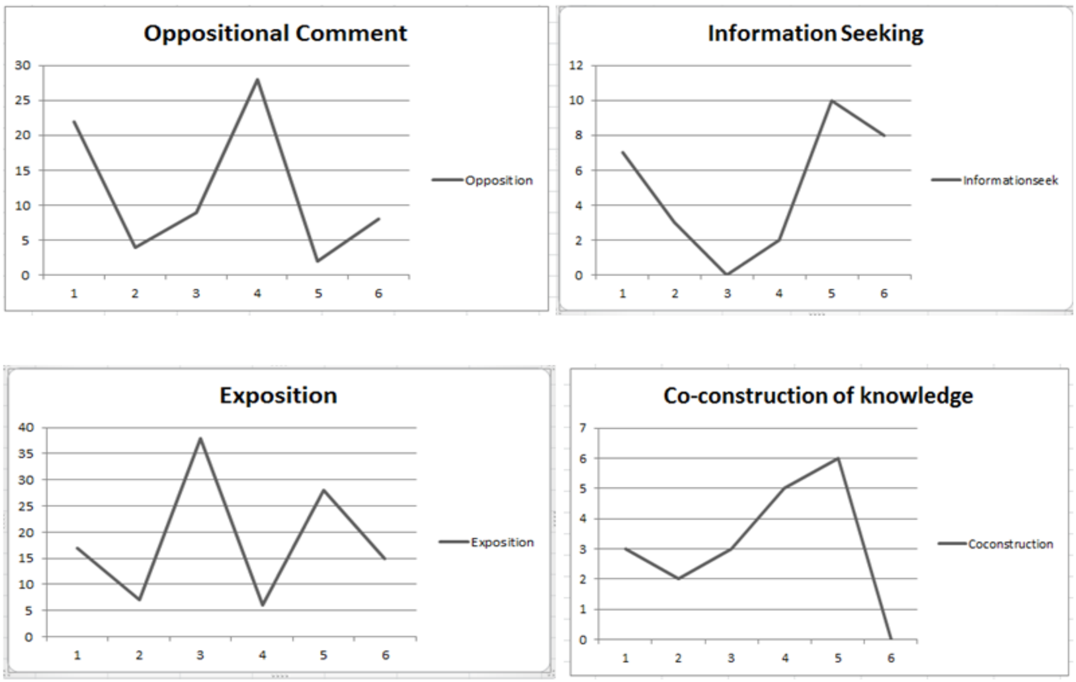



*Figure 4. 16* Weekly changes of engagement components for Class 2

Figure 4.16 suggested that none of the four engagement components increased throughout the study. All of the four engagement components fluctuated from week to another week. For example; oppositional comment was used at most in week 4 (28 times). Likewise, class 2 students were highly active in using opposition comments in week 1 (22 times), but they used oppositional comment very few in week 5 (2 times) and week 2 (4 times).

Regarding information seeking graph, Figure 4.16 shows that class 2 students generally used information seeking in week 1 (7 times), week 5 (10 times) and week

6 (8 times). On the other hand, students rarely used information seeking in week 2 (3 times) and week 4 (2 times), and information seeking was not used in week 3.

Another engagement component is expositional comment and students used exposition comment at most in week 3 (38 times). This component was also frequently used in week 5 (28 times). On the other hand, expositional comment was used least in week 4 (6 times) and week 2 (7 times).

Co-construction of knowledge was the least used component in class 2 similar to class 1. For example, this component was hardly ever used in first three weeks (2-3 times) and co-construction of knowledge was not used in week 6. On the other hand, this component was used relatively more in week 4 (5 times) and week 5 (6 times).

In sum, most used engagement component in class 2 was expositional comments in general. Oppositional comments followed expositional comments. On the other hand, information seeking and co-construction of knowledge were less frequently used in class 2. There was no obvious increase or decrease in frequency of four engagement components in class 2. Comparison of class 1 and class 2 also suggested that class 2 students used engagement components more than class 1 students in general. Therefore, it can be said that class 2 students were more active than class 1 students regarding engagement in argumentation process.

After findings of class 1 and class 2 are presented in this part, assertions derived from similarities between two classes are explained in terms of students' engagement in argumentation.

### 4.5.3. Assertions derived from similarities between class 1 and class 2.

Two classes who participated in qualitative part of the study had some commonalities. These commonalities allow claiming assertions regarding the nature of students' engagement in argumentation process in middle school level. Accordingly, seven different assertions have been claimed. Three of these assertions are related with the nature of activity (i.e., content of the week). Another four assertions are specific to components forming students' engagement in argumentation process (e.g., expositional comment). At first, assertions about nature of activity are presented. After that, assertions specific to components forming students' engagement in argumentation process are given.

#### *4.5.3.1. Assertions about nature of activity.*

Similarities between class 1 and class 2 revealed three assertions about nature of activity. In this study, Turkish middle school science curriculum (MONE, 2013) was followed. Activities were prepared based on two units (matter and heat and electricity) found in curriculum. Accordingly, different curricular objectives were followed in each week. Therefore, content and activities of the weeks were different from each other. It is thought that differences of content and activities between weeks had effect on students' engagement in argumentation process. Three assertions about nature of activities and their impact on students' engagement in argumentation support this idea. The assertions are as follow:

Assertion 1: Students engage in argumentation at most when evidence cards are provided to them.

Findings of the study suggested that both classes engaged in argumentation at most in week 3. While class 1 engaged in argumentation 59 times in this week, class 2 engaged in argumentation 50 times. Content of the week 3 was thermal insulating products. In this week, students did not conduct experiment. On the other hand, evidence cards including definition of thermal insulating products and their characteristics were shared with students. By using evidences, students constructed their arguments and engaged in argumentation. Although students engaged in argumentation at most in week 3, their engagement mostly included expositional comments in this week for both class. In their expositional comments, students proposed their ideas when matching thermal insulating product and the place where it should be used in the house. While doing this, students justified their ideas using evidence cards. A lot of students participated in argumentation process and one of the examples occurred in class 1 in week 3 is presented below:

Student 134: I think that plastic foam should be used in interior wall because it keeps house warm and there is no heat loss (Expositional comment)

Student 122: Plastic foam is durable so it can be used in interior wall (Expositional comment)

Student 143: There is a space in plastic foam and this space provides heat insulation, therefore, heat does not lose (Expositional comment)

As it is seen, students constructed their arguments using the information found in evidence cards and they proposed their ideas. Although expositional comments dominated week 3, students sometimes used oppositional comments in this week for both classes. The following excerpts are examples of oppositional comments that both classes' students used in week 3:

Class 1 instance:

Student 131: I think that wood wool should not be used in interior wall because it is too expensive; its cost is 180 liras. Moreover, it is not durable and it can be eaten by insects... (Expositional comment)

Student 133: You said that insects can harm to the wood wool, however, if you add chemicals on wood wool, insects cannot approach to the wood wool and it does not get harm (Oppositional comment)

Student 111: I want to say something to you (refers student 133). If we use wood wool, this shows that we are using environmentally friendly products. If we are environmentally friendly, why do we use chemicals to protect wood wool? Chemicals are not environmentally friendly (Oppositional comment).

Student 133: But, chemicals used to protect wood wool may not harm to environment (oppositional comment).

Class 2 instance:

Student 243: I think that tar can be used in ceiling as insulating material because tar's color is dark. If it is used for ceiling, dark color absorbs heat and tar may adjust room temperature (Expositional comment).

Student 231: I think tar should not be used. You claimed that dark color absorb heat. In evidence cards, it is written that tar is flammable material. If it absorbs heat as you said, then it flames and home flames too (Oppositional comment).

Similar to week 3, both classes were active in week 4. Both classes engaged in argumentation 41 times in this week. Content of the week 4 was types of fuels. The common point between week 3 and week 4 was not conducting experiments and using evidence cards to construct argument, therefore, it can be summarized that 6$^{th}$ grade students engaged in argumentation more when they did not conduct experiment and they use already available data exist in evidence cards.

Assertion 2: Students engage in argumentation less in a following week when they do similar activities.

Findings showed that both classes engaged in argumentation least in week 2. Accordingly, class 1 students engaged in argumentation 12 times in week 2 and class 1 students engaged in argumentation 16 times. This situation might be related with task similarity between week 1 and week 2. First week core idea was heat conductors and students tried to find which kind of spoons' heat conductivity is better than others in their experiments. Students engaged in the whole class discussions in this week to some extent. For example, class 2 students were highly active in this week. They engaged in argumentation 49 times in first week. Besides, in second week core idea was heat insulators and students examined which kind of cup is a better heat insulator. This week activity was similar to the previous week's activity because both heat conductors and heat insulators belong to same main idea which is heat transfer. Likewise, students used similar materials in their experiments. For example, they used hot water, cup, thermometers and easily accessible materials such as aluminium foil in both weeks' experiments. Therefore, students might think that they do similar things in first two weeks. The idea of repeating same activities might cause students to lose their excitement in week 2 and they engaged less in argumentation. This claim can be true especially for class 2 because while class 2 students engaged in argumentation 49 times in week 1, they engaged in argumentation 16 times in week 2. On the other hand, this claim can be less true for class 1 because class 1 was not active in week 1 and week 2, but still it can be claimed that class 1 students' engagement decreased from week 1 (19 times) to week 2 (12 times) depending on the total number of their engagement in argumentation. The followings are the examples of students' engagement in argumentation in week 2:

Class 1:

Student 111: In this experiment, our research question was which cup does lose its heat faster. We hypothesized that metal cup loses its heat faster... The temperature difference between metal cup's water initial temperature and final temperature is 17 degree. On the other hand, temperature difference for plastic cup's water is 19.5 degree. This means plastic lost its heat faster (Expositional comment).

Student 143: I think that metal cup transfers its heat faster to the surface where it contacts because it is better conductor, but plastic cup do not lose its heat as fast as metal cup. Therefore, I think that metal should lose its heat faster than plastic in your experiment (oppositional comment).

Student 141: When we poured hot water to the metal cup, metal cup might be cold and there can be heat transfer between hot water and metal cup before we measured initial temperature of metal cup. By this way, we might measure the initial temperature of water less than actual value, so heat difference might be less than actual heat difference (co-construction of knowledge).

Class 2:

Student 223: In this experiment, we examined which cup transfer heat more. We hypothesized that metal cup transfers more... We measured temperature of hot water in 1. , 5. and 10. minutes using thermometer. While metal cup's water lost 22°

temperature, glass cup lost 25˚. Maybe our result is wrong, but we found these values (Expositional comment).

Student 243: Maybe you touched the thermometer too much and this might cause temperature increase on value you read for metal cup (co-construction of knowledge).

Assertion 3: Use of ideal conditions might inhibit students' engagement in argumentation.

Students conducted experiments about the factors affecting bulb brightness in the last week (i.e., week 6). However, their experiment results were faulty and they could not reach the correct explanations in the pilot study. Therefore, researcher removed real experiments from the study, and found a simulation program for the main study. By this way, students could conduct experiments and reached the scientifically acceptable conclusions. Although students' conclusions were acceptable regarding the factors affecting bulb brightness, students' engagement in argumentation process decreased for both class compared with the pilot study. The possible reason explaining why students engaged in argumentation less when they used simulation can be related with idealized conditions. When students conducted experiments in laboratory, they reached faulty results, they could not conduct experiment, but they formulated their idea based on their observations. These ideas were different from each other and these diverging points might let them discuss more in the pilot study. On the other hand, when students conducted their experiments through simulation programs, all conditions were idealized and there was no error in experiments. Besides, students mainly reached similar results. These idealized conditions might clear all question marks in students mind, and so students might not discuss further in this week. The following records are examples of students' whole class discussions which were done after students completed their experiments using simulations:

Class 1:

Student 121: Our research question is “Does thickness of wire in circuit affect bulb brightness?” We hypothesized that it affects brightness...We connected circuit components and measured the current. Then, we increased the thickness of wire, and re-measured the current. According to our findings, current was 0.35 when wire was thin. After we increased wire thickness, current increased to 0.46. Our claim is when we used thick wire, bulb brightness increased (Expositional comment)

Student 131: I have no question about circuit, but what was the ampere when you used thick wire, was it 0.45? (Information seeking)

Student 123: 0.46 (Exposition)

Student 121: When we used thin wire, we read 0.35 (Exposition)

Class 2:

Student 223: Our research question was “Does the length of a wire affect light brightness?” We hypothesized that it does not affect, but this hypothesis did not happen... When we increased the length of wire, we observed that bulb brightness is affected by change of wire length. When we increased the wire length, bulb lights less (Exposition)

Student 234: Why did you think that wire length does not affect bulb brightness prior to experiment? (Information seeking)

Student 223: No matter how long you increase the length of wire, you have same electricity, so we thought bulb brightness does not change (Exposition)

Student 234: But it changed (Oppositional comment)

Student 223: Yes, it changed (Consensus-No further discussion).

As it is seen from Figure 4.11 and Figure 4.14, there is no obvious sign that students’ engagement in argumentation develop over time. One factor explaining why students did not engage in argumentation more over time can be nature of activities given above. For example, if students prefer to discuss more when they get evidence cards, it is expectable that students will not discuss so much when they conduct experiment or they do not use evidence cards in following weeks. Similarly, it is not meaningful to expect students to discuss more when current activities are similar to previous weeks' activities. Likewise, students may not engage in argumentation so much when the conditions are idealized as in the week when students used simulation programs to conduct their experiments. As a result, nature of activity might affect development of students' engagement in argumentation process. Up to now, assertions about nature of activities derived from commonalities of two classes were presented. Assertions specific to components forming students' engagement in argumentation process derived from similarities of two classes are explained in next part.

#### *4.5.3.2. Assertions specific to components forming students' engagement in argumentation process.*

Four components which are expositional comment, oppositional comment, information seeking, and co-construction of knowledge were obtained from Sampson and Clark's (2011) work to explain nature of students' engagement in argumentation process. Analysis of two classes' argumentation process revealed four commonalities regarding components forming students' engagement in argumentation. These four commonalities about components (e.g., oppositional comment) are presented as assertions in this part.

Assertion 1: When students engaged in argumentation, they mostly use expositional comment.

Findings of two classes showed that students usually use expositional comment when they engage in argumentation. Through expositions, students propose, justify, and clarify their ideas. By this way, they constructed their arguments. Apart from constructing arguments, expositional comments might be seen as starting point of argumentation process because students propose ideas through expositions and other students added their ideas based on these expositional comments. These added ideas can be oppositional comments, information seeking, co-construction of knowledge or expositional comments. However, dominating engagement component in this argumentation process was expositional comments for both classes. An example of proposing expositional comment and adding ideas on this component sustaining argumentation process was given in following example. The argumentation resulted in agreement in this example.

Class 2:

Student 234: We think that we should use solar energy because it is renewable. We do not want to use wood because it is non-renewable (Exposition).

Student 231: I want to add something. We should not use nuclear energy too because it is risky. Actually, nuclear energy is good for economy, but it damages to environment (co-construction of knowledge).

Student 223: You said that nuclear energy damages to environment, but it also damages to out of world. Nuclear energy has radiation; this radiation can pass to space and damage the space (co-construction of knowledge).

Student 231: However, space already has radiation in it. The radiation that space has is much more than nuclear power plants had (oppositional comment).

Student 223: Yes, you can be right (Agreement).

On the other hand, study showed that students did not use expositional comments so many in week 2. It is possible that students do not propose their ideas when activity is similar to previous week's activity because they might lose their excitement. The loss of excitement might cause not to proposing their ideas as expositional comment, and not proposing their ideas might inhibit revealing other three components of engagement in argumentation process. To sum up, students engaged in argumentation least in week 2 whose activity was highly similar to week 1.

Although students did not produce expositional comments so many in week 2, both classes used expositional comments with the highest frequencies in week 3. The core idea of week three was thermal insulation products and students had no prior knowledge on this topic. Therefore, they just depended on the evidence cards provided them while constructing arguments. According to this, students might propose expositional comments at most when they have no prior knowledge and already available data is presented to them. While students use expositional comments with the highest frequencies using the available data, they might focus on their own ideas ignoring others. In other word, focusing on their ideas and ignoring alternative views might inhibit the occurrence of other three components forming students' engagement in argumentation. Following is an example that students focused on constructing their arguments and ignored other students' ideas in week 3 in which they had no prior knowledge and they used evidence cards.

Class 1:

Student133: I think that glass wool should be used in interior wall because it keeps warm the house (Expositional comment)

Student 132: Glass wool is not flammable, it is economical and durable, and so I can use it for interior wall (Expositional comment)

Student 122: It can be used every part of the house, so I can use glass wool in interior wall (Expositional comment)

Student 134: Glass wool does not decay; it protects itself from the insects, so I can use glass wool in interior wall (Expositional comment)

As the excerpts given above show, students might focus on their ideas and they just propose expositional comments and do not consider alternative ideas in some situations. This may not be acceptable for argumentation because students do not use other components forming their engagement in argumentation in such situations. This was observed in this study when already available data was given to students in a topic that students had no prior knowledge. Therefore, it might be better to avoid giving students already available data when they had no prior knowledge and alternative plans for effective argumentation might be prepared.

Assertion 2: If evidence cards are provided to students and students have prior knowledge about the topic, they mostly tended to use oppositional comments in argumentation process.

Oppositional comments are the challenging the ideas of others (Sampson & Clark, 2011). In this study, both of the classes used oppositional comment at most in week 4 which is about type of energy sources. Both classes were highly active in pre-discussions which was done prior to the main activity in week 4, therefore, it can be understood that students had prior knowledge about types of energy sources. Similar to week 3, evidence cards were given to students in week 4. Evidence cards included information about definition, advantages and disadvantages of energy sources. According to Figure 4.12, 39 % of the whole class discussions were oppositional-comments in class 1. Similarly, Figure 4.15 shows that 68 % of class 2 students' engagement in argumentation corresponded to oppositional comments. The following excerpts are students' oppositional comments used in week 4:

Class 1:

Student144: I think that we should use solar panels because it is natural and it does not damage to environment. It has many advantages. On the other hand, coal should not be used as energy sources because it contaminates the environment (Expositional comment).

Student 111: You want to use solar energy; however, solar energy is not sufficient for world demands. I want to add something about coal. We can easily find coal, so it can be used as energy sources. Moreover, coal provides more heat comparing with solar energy (oppositional comment).

Student143: I will answer about solar energy part. If we place solar panels to the deserts and educate people, I think solar energy supplies all world's demands (oppositional comment).

Student 111: Ok, let's place them in Egypt, but we are in Turkey. We cannot benefit from the solar energy for our demands (oppositional comment).

Class 2:

Student 244: I think that we should use wind energy because it is free. It does not cause global warming, and it is renewable. I think we should not use wood as energy source because it is non-renewable (Expositional comment).

Student 223: However, solar energy is free too. We prefer to use solar energy. Solar energy is better energy comparing with wind energy because wind does not always available, but sun rises every day. If there was no sun, there would be no light outside. Therefore, solar energy should be used (Oppositional comment).

Student 231: You said that wind does not always available, but sun is also not available at nights (Oppositional comment).

Student 221: However, earth rotates. When somewhere on earth is day, another place is night. Hence, sun is always available (Oppositional comment).

Student 231: However, you collect energy from just one point, not from different parts of the earth (Oppositional comment).

Student 221: We know when the sun rises and sets, but we do not know when wind blows (Oppositional comment).

Student 231: I say that wind energy close the gap. Wind provides the same energy that solar energy provides us (Oppositional comment).

Student 223: Solar energy is always available (Oppositional comment).

Student 231: You can create artificial wind and close the gap (Oppositional comment).

Student 223: We do not talk about artificial wind (Ignore).

Although oppositional comment was the most used component forming students' engagement in argumentation in week 4, it was rarely used in week 2 for both classes. While class 1 used oppositional comment one time in week 2, class 2 used it four times in week 2. Why students did not frequently produce oppositional comments in this week might be also related with loss of excitement. As it was explained in the previous assertion, it is possible that students might think that week 2 activity was so similar to week 1 and so they might lose their excitement. On that account, students might not engage in argumentation and use components forming their engagement in argumentation. Therefore, not using similar activities in following weeks of an argumentation study can be better if students are expected to use all engagement components including oppositional comments.

Assertion 3: Students use information seeking component more in the weeks when they conduct experiment compared with weeks when they do not conduct experiment. Moreover, use of information seeking becomes top level when students conduct experiment at the beginning of the unit.

Information seeking component of students' engagement to argumentation refers requesting more information and clarification (Sampson & Clark, 2011). One of the commonalities between classes showed that 6$^{th}$ grade students mainly used information seeking component in first weeks of the units. According to Figure 4.12, 32 % of class 1 discussions were information seeking at the beginning of the matter and heat unit (i.e., week 1). Likewise, another week that class 1 students actively used information seeking was first week of the electricity unit (i.e. week 5). 14 % of total discussion was information seeking in this week. Same trend was also observed in class 2. According to Figure 4.15, class 2 students actively used information seeking at the beginning of both units. Percentages of information seeking used in class 2 for week 1 and week 5 were 14 % and 22 % respectively. Some examples that students used information seeking in these weeks are presented below:

Class 1:

Student143: In this experiment we examined which object's heat conductivity is more. We hypothesized that metal is better conductor and wood is better insulator... After we removed spoons from hot water, we touched each object and tried to understand which one is hotter. Our observations showed that initial temperature of water was 70˚. Metal spoon got hot in few minutes. It got 2 minutes, but plastic spoon got heat in four minutes. Wood spoon wetted because of water. It became heavier because it absorbed some water in cup... When experiment ended, water temperature was 41˚ (Exposition, week 1)

Student 111: What was the temperature of spoons? (Information seeking)

Class 2

Student 223: Our research question was "Which kind of cup does conduct electricity better?"... After we completed the circuit, we connected the different cups in turn. When we connected the cups, we observed the brightness of bulb. We repeated this procedure for each cup. The cups that we used were foam, metal and plastic. There was little brightness when we connected metal cup; however, bulb did not light when we connected other two cups. Bulb brightness was at most when we completed the circuit without connecting anyone of the cups. Our claim is metal cup conducts electricity better because bulb lighted when we connected metal cup (Exposition, week 5)

Student 234: How much did bulb light? Because you said there was little brightness (Information seeking)

Student 223: When we completed the circuit at the beginning, it was so bright. When we connected metal cup, the brightness of bulb was less comparing with previous condition (Exposition)

Student 231: Why did metal conduct electricity? (Information seeking)

As it was seen from the given excerpts; students mainly used information seeking in three different conditions. First, students use information seeking when they have no idea about the procedure of group who conducted the experiments. For example; class 1 students asked the temperature of spoons to understand which spoon was hotter. Second, students used information seeking when there was uncertainty. For example; there was uncertainty about the brightness of the bulb in class 2 excerpt. Therefore, class 2 students used information seeking to be knowledgeable about the amount of brightness. Third, students tend to use theoretical questions in their information seeking to understand the reason of observed phenomena. For instance, students presented the experiment results in class 2 excerpt. However, other students did not satisfy just listening the electrical conductivity results and they used information seeking to understand why some objects conduct electricity better than others.

Although students tended to use information seeking when they conducted experiments, especially at the beginning of the units, both classes hardly ever used information seeking in the weeks that they did not conduct experiment or evidence cards were presented to them. According to Figure 4.12, class 1 students' use of information seeking corresponded to 2 % and 12 % of total discussions in week 3 and week 4 when evidence cards were given to them. Similarly, class 2 students never used information seeking in week 3 and percentage of information seeking was 5 % of week 4's whole class discussions.

In conclusion, this data may suggest that use of information seeking is top when students had no information about other groups' experiment and theoretical content of topic. Therefore, students might use information seeking most at the beginning of the weeks that they conducted experiments. In these weeks, students asked both procedural and theoretical questions. On the other hand, students might think that they have necessary data to construct argument in the weeks when evidence cards were provided them. Hence, students might not focus on information seeking in these weeks.

Information seeking might have specific role in students' engagement in argumentation like expositional comment. In assertion 1, it was explained that argumentation process starts with expositional comments in which students propose their ideas. However, expositional comments are not sufficient to maintain argumentation. For example, both classes used expositional comments so many times in week 3; however, other three components forming students' engagement in argumentation were not used in this week. In this point, information seeking might have key point to sustain argumentation. Accordingly, students might use expositional comments as data sources to engage in argumentation. After presenting group tell their expositional comments, other groups of students might identify missing points, conflictions and uncertainties of expositional comments. By this way, they might start to examine expositional comments through information seeking. As new questions (i.e., information seeking) and explanations (i.e., expositional comments) emerges, diverging or converging points between different ideas might emerge too which decide the direction of argumentation.

Assertion 4: Students usually do not use co-construction of knowledge in argumentation.

Co-construction of knowledge is summarizing, revising, supporting and adding ideas to others' ideas (Sampson & Clark, 2011). According to Figure 4.11 and Figure 4.14, both classes used co-construction of knowledge less than other three components forming students' engagement in argumentation. Moreover, frequency of number of co-construction of knowledge decreased further for both classes when similar activities were used in following weeks (i.e., week 2) and when simulation activity was used (i.e., week 6). The following excerpts are examples of co-construction of knowledge that students used in whole class discussions.

Class 1

Student 111: Our research question is "What is the effect of aluminium foil on the rate of melting butter?" (Expositional comment)

Student 132: I think you can try two conditions to test this experiment. In first condition, you can wrap butter with aluminium foil and throw it in hot water. In second condition, you throw the butter in hot water without aluminium foil. Then, compare the rate of melting of butter for two conditions (co-construction of knowledge)

Student 111: ok, that sounds fine; we can use your suggestion (Accept).

Class 2

Student 213: Our research question is “Which kind of water does conduct electricity?”... Our hypothesis was salty water conducts electricity and the result was as we expected. Bulb lighted when we add salty water to the circuit, but bulb did not light when we add other kinds of water like sugared water (Exposition)

Student 215: Water became yellowish a few times later (Exposition)

Student 234: Sugared water did not conduct electricity in our experiment too (Co-construction of knowledge)

Student 224: Why did water color become yellowish? (Information seeking)

Student 213: I do not know.

As it is seen, class 1 excerpt is evidence that students use co-construction of knowledge when the presenter group has difficulty in conducting experiment. In this excerpt, one of the students suggested an experimental design by adding his ideas (i.e. co-construction of knowledge) and presenter group accepted his ideas and chose it as data collection procedure. Similarly, students use co-construction of knowledge to support others' ideas. For example in class 2 excerpt, one student encouraged presenter group saying that they reached the same result regarding salty water's electrical conductivity. Moreover, co-construction of knowledge was used when students did not have enough information to explain the natural phenomena. In such situations, students tried to reach correct explanations by sharing and adding ideas of each other. The following is an example that students collaborate with each other to explain natural phenomena. Although students reached wrong explanation, they still put effort to reach an explanation by using co-construction of knowledge.

Class 2:

Student 245: In our experiment, foam and paper cup did not conduct electricity (Exposition)

Student 223: Do you know the reason why foam and paper cup did not conduct electricity? (Information seeking)

Student 241: because they are insulator (exposition)

Student 231: I think there is a space between foam particles and so electricity does not pass through it (exposition)

Student 241: This was same in heat insulation. The reason of insulation was the particular structure of matter. Particular structure of matter can be cause of electrical insulation too (co-construction of knowledge).

Student 234: Yes, foam is electrical insulator because of the space between particles. This space does not allow electricity to transfer from somewhere to another (co-construction of knowledge).

Even though reaching scientifically wrong explanation is seen as an undesired situation in terms of learning at first glance, such kind of wrong explanation might be source for further discussion. For example, scientifically wrong explanation can be examined by students and teacher guides the process. Examination of scientifically wrong explanation might trigger further discussion, students engage in argumentation more and their conceptual understanding might increase. As a result, students might reach the scientific explanation.

As it was explained above, students used co-construction of knowledge less than other three components forming students' engagement in argumentation throughout the study. However, number of frequency of using co-construction of knowledge decreased further when activities were similar to each other. It is possible that students might think week 2 activity was similar to week 1, therefore they did not engage in argumentation, and so they did not produce co-construction of knowledge in this week. Similar trend was also observed in last week when students used simulation to examine the factors affecting light brightness. Accordingly, idealized conditions revealing in simulation activity might negatively affect students' engagement in argumentation including co-construction of knowledge component.

In this part, qualitative findings of the study were examined in terms of students' engagement in argumentation process. According to the results, both classes had some commonalities. Using these commonalities, two different types of assertions were proposed which are assertions about nature of activities and assertions specific to components forming students' engagement in argumentation process. Assertions about nature of activities included three assertions. Activities about nature of activities firstly showed that students engage in argumentation at most when evidence cards are provided to them. Second assertion about nature of activities indicated that students engage in argumentation less in following week when they do similar activities. Last assertion about nature of activities suggests that idealized conditions might inhibit students' engagement in argumentation. On the other hand, assertions specific to components forming students' engagement in argumentation process revealed four assertions. These assertions were unique to components (e.g. oppositional comment) that form students' engagement in argumentation process. Assertion about expositional comment suggests that when students engaged in argumentation, they mostly use expositional comment. Another component forming students' engagement in argumentation is oppositional-comment and assertion about this component suggests that students use oppositional comments at most when they have prior knowledge about topic and evidence cards are provided them. Regarding information seeking component, assertion shows that students tend to use information seeking when they conducted experiments, especially at the beginning of the units, however students hardly ever use information seeking in the weeks that they do not conduct experiment. Finally, assertion about co-construction of knowledge demonstrates that students usually do not use co-construction of knowledge in argumentation. As findings of both classes had some commonalities, these two classes' findings also included some differences. These differences between classes allow proposing further assertions explaining the nature of students' engagement in argumentation. The assertions derived from differences between two classes' findings are presented in next part.

#### 4.5.4. Assertion derived from differences between class 1 and class 2.

Although argument-based inquiry was implemented in both classes, whole class discussions included some differences as seen from Figure 4.11 and Figure 4.14.

Comparison of Figure 4.11 and Figure 4.14 shows that class 2 students engaged in argumentation more than class 1 students. Therefore, class 2 can be labelled as more active class and class 1 can be labelled as less active class. Accordingly, one assertion was proposed based on the comparison of more active class and less active class. It is thought that assertion derived from differences of two classes further enlighten nature of students’ engagement in argumentation process.

Assertion: While more active class tend to engage in argumentation process more in the weeks when they conduct experiments, less active class tend to engage in argumentation process more in the weeks when students use evidence cards to construct arguments.

Data was further analyzed based on percentages of each component considering different units and sources of data. Accordingly, first two weeks were about matter and heat unit and students conducted experiments and used first hand data. Week 3 and week 4 were about matter and heat topic again however, evidence cards were given to students in these weeks and students did not conduct experiment. They used second hand data found in evidence cards to construct arguments in these weeks. Last two week was about electricity unit and students conducted experiments and used first hand data they obtained from experiments. In this analysis, total frequencies of each component for related weeks were divided to overall frequencies of that component for each class. For example, class 1 used expositional comments 101 times throughout the study. This class used expositional comment 8 times in first week and 6 times in second week, so class 1 used expositional comments 14 times for first two weeks that students conducted experiment in matter and heat unit. Total frequency of these two weeks which is 14 was divided to overall frequency that is 101. By this way percentage of first two weeks’ expositional comment for class 1 was calculated as 14 %. Same calculations were done for each components, each classes and corresponding weeks. Percentages of each component forming students’ engagement in argumentation for each class considering units and sources of data are presented in Table 4.20;

Table 4. 20

*Percentages of Each Engagement Components regarding Units and Sources of Data*

| | First two weeks: Matter and Heat Unit First hand data | | | | Third and Forth weeks: Matter and Heat Unit Second hand data | | | | Last two weeks: Electricity Unit First hand data | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Expositional Comment | Oppositional Comment | Information Seeking | Co-construction of knowledge | Expositional Comment | Oppositional Comment | Information Seeking | Co-construction of knowledge | Expositional Comment | Oppositional Comment | Information Seeking | Co-construction of knowledge |
| More Active Class | 22% | 37% | 33% | 26% | 40% | 52% | 6% | 42% | 39% | 14% | 60% | 32% |
| Less Active Class | 13% | 4% | 31% | 38% | 58% | 57% | 27% | 29% | 28% | 39% | 40% | 33% |

Both classes used co-construction of knowledge seldom throughout the study. Therefore, co-construction of knowledge considering their percentages was not compared between classes. According to Table 4.20, percentages of three components which are expositional comment, oppositional comment, and information seeking were higher in more active class than less active class for first two weeks when students conducted experiment in matter and heat unit. While percentages of expositional comments, oppositional comments and information seeking occurred in more active class's first two weeks were 22%, 37% and %33 respectively, percentages of these components used in less active class in first two weeks were 13 %, 4 %, and 31 %. Likewise, percentages of expositional comment and information seeking occurred in more active class was higher than less active class in last two weeks that students conducted experiments in electricity unit. While percentage of expositional comment in more active class was 39 %, it was 28 % in less active class. Similarly, percentage of information seeking was 60 % in more active class, but it was 40 % in less active class. On the other hand, percentages of oppositional comment in less active class outperformed this component's percentage in more active class for last two weeks (14 % vs. 39 %). This superiority of less active class regarding oppositional comments towards more active class in last two weeks might be an exception.

To sum up, this part focused on the differences between two classes during argumentation. Differences of the students suggested that class 2 engaged in argumentation more than class 1. Differences between two classes allowed proposing one assertion. Accordingly, while more active class tend to make argumentation in the weeks when they conduct experiments, less active class tend to engage in argumentation in the weeks that students use evidence cards to construct arguments. It is thought that both assertions derived from similarities and differences between classes contributes on enlightening $6^{th}$ grade students' nature of engagement in argumentation process.

### 4.6. Summary of Results

In conclusion, quantitative findings of the study suggested that argument based inquiry treatment improved $6^{th}$ grade level students' content knowledge about matter and heat topic and electricity topic. On the other hand, students' epistemological beliefs and science process skills did not improve.

This study also sought to explain argumentation schemes that $6^{th}$ grade students used in argument based inquiry treatment. Similarities and differences of two classes were used to claim assertions that enlighten middle school students' argumentation schemes in ABI treatment. Similarities between classes showed that total number of types of argumentation schemes increased over time for both classes. Similarities also showed that students tend to use some argumentation schemes more than others in science classes. These argumentation schemes that both classes' students mainly use are argument from expert opinion, argument from sign, argument from position to know, argument from comparison, and argument from evidence to hypothesis. Findings also suggested that types of data (i.e. first hand data, second had data) might affect the argumentation schemes students used. Accordingly, students might use argument from evidence to hypothesis when they use first hand data, but they use argument from expert opinion when they use second hand data. Differences between two classes were also beneficial to understand nature of students' argumentation schemes. According to differences emerged between two classes, this study showed that more active class (i.e. Class 2) uses more argumentation schemes comparing with less active class. This trend was obvious especially after week 3. Differences also showed that while most used argumentation scheme in more active class was argument from position to know, the most used argumentation scheme in less active class was argument from expert opinion.

Findings also informed about students' engagement in argumentation process. Two classes' video records were compared with each other and assertions were identified based on the similarities and differences of classes regarding their engagement in argumentation process. Similarities of two classes showed that both classes engaged in argumentation most in the weeks when they used evidence cards. On the other hand, students appeared to engage in argumentation less if following week's activity is similar to previous activities. Likewise, use of ideal conditions instead of real experiments might inhibit students' engagement in argumentation process. Similarities of classes also showed that the most used engagement component was expositional comments which occur when students propose and clarify their ideas. Findings also suggested that both classes used oppositional comments most when they have prior knowledge, and evidence cards are given him. Next, students usually used information seeking component of engagement at the beginning of the units; moreover, students tend not to use this engagement component in the weeks when they do not conduct experiments. This study also showed that both classes usually did not use co-construction of knowledge in argumentation.

Classes also had some differences when they engaged in argumentation. Accordingly, class 2 engaged in argumentation more than class1. Therefore, class 2 was labelled as more active class and class 1 was labelled as less active class in this study. Findings emerging from differences of two classes showed that while more active class tend to engage in argumentation in the weeks that students conduct experiment, less active class tend to engage in argumentation in the weeks that evidence cards were provided to students.

# CHAPTER 5

# DISCUSSION

This study attempted to investigate effect of ABI treatment on 6$^{th}$ grade students' scientific literacy components which are content knowledge, epistemological beliefs, and science process skills. Moreover, this study uncovered participants' argumentation schemes and their engagement in argumentation process. In line with this, effects of ABI treatment on scientific literacy components are discussed at first. Second, the findings about argumentation schemes and then students' engagement in argumentation process are discussed. At the end, implications of the study are presented.

## 5.1. Argument Based Inquiry and Scientific Literacy

In this study, vision-1 for scientific literacy was adopted. Accordingly, students were expected to act like scientists and they were expected to focus on both process of science and product of science (Roberts, 2007). Therefore, variables which are directly linked with scientific enterprise (e.g., content knowledge, epistemological beliefs, and science process skills) were selected and it was assumed that these variables represent scientific literacy. Accordingly, ABI's effect on students' content knowledge, epistemological beliefs and science process skills are discussed in this part. It is expected that these discussions about content knowledge, epistemological beliefs and science process skills shed light the effect of ABI on students' scientific literacy.

### 5.1.1. Argument Based Inquiry and Content Knowledge.

In this study, content knowledge is defined as academic content of a discipline (e.g., physics, biology, chemistry) (Carlson & Daehler, 2019). Students' conceptual understanding about matter and heat unit and electricity unit taught in $6^{th}$ grade level represents their content knowledge. In the present study, the effect of ABI treatment on sixth grade students' content knowledge in Matter and Heat and Electricity topics was examined using multiple choice tests. The test were developed in the current study and demonstrated to provide valid and reliable measure of students' content knowledge in these topics. In order to investigate the effectiveness of the ABI treatment, Within Subject Repeated MANOVA was conducted. More specifically, this analysis was carried out to explore the changes in participants' content knowledge about matter and heat topic and electricity topic in ABI treatment. Results showed that there was a significant increase in the level of students' content knowledge in both topics with a large effect size. At this point, it is also important to note that, the effect size was larger for the second topic (i.e., electricity unit) suggesting that the effectiveness of the ABI treatment on content knowledge increased over time. However, there was no control group in this study which is one limitation in quantitative part. Therefore, it is not certain to what extent ABI treatment was effective to improve participants' content knowledge compared to regular instruction. On the other hand, previous ABI research mainly employed experimental research including control groups. These previous ABI treatments mainly suggested that experimental groups outscored control groups (Erkol et al., 2017; Greenbowe et al., 2007; Hand et al., 2004; Hohenshell & Hand, 2006; Kingir et al., 2012; Kingir et al., 2013; Shamuganathan & Karpudewan, 2017; Taylor et al., 2018; Yeşilçağ-Hasançebi & Günel, 2013) although there are some exceptions showing no statistical difference between experiment and control groups (Cronje et al., 2013; Rudd et al., 2001).

Why participants' content knowledge increased in this treatment is one focus of interest. Previous research mainly explained contribution of ABI treatment on content knowledge referring to ABI's focus on process of science (Kingir et al., 2013), the relationship between ABI and conceptual change (Kingir et al., 2013), use of SWH as scaffold (Akkus et al., 2007; Chen et al., 2016; Cronje et al., 2013; Hand et al., 2004), group discussions (Hand et al., 2004; Heng et al., 2015; Shamuganathan & Karpudewan, 2017), following ABI phases strictly (Cronje et al., 2013; Erkol et al., 2017), use of students own research questions and allowing them to follow their own data collection process (Cronje et al., 2013; Hand et al., 2004).

Concerning the effect of ABI treatment on content knowledge through its focus on the process of science, it is pointed out that students in ABI classes act like scientists, they ask questions, conduct investigations, interpret data and discuss on different interpretations in this process. As a result, students construct knowledge in this process and their content knowledge increases (Kingir et al., 2013). Similarly, students, in the present study, acted like scientists and they focused on the process of science especially in the weeks when they conducted their own investigations (week 1, week 2, week 5 and week 6). Therefore, their engagement in process of science might have improved their content knowledge in this study.

In addition, Kingir et al. (2013) claimed that ABI is consistent with conceptual change approach and so it facilitates students' conceptual understanding. According to this, students get surprised when they obtain conflicting data in data collection process and they become suspicious to their own prior knowledge. These conflicting data and suspicion on their prior knowledge provoke students and they start to discuss on conflicting data and their prior knowledge. These discussions end with negotiations. By this way, students learn from each other (Kingir et al., 2013). The consistency between conceptual change and ABI was also observed in this study. For example, one of the groups hypothesized that length of wire does not affect wire resistance because the voltage and current should be same. On the other hand, the group conducted the experiment and observed that there is a reverse proportion between wire length and wire resistance. This inconsistency led group to engage in discussion. Other groups used an analogy to explain this reverse proportion. By listening other groups' analogy, the group learnt the reason of reverse proportion and their content knowledge increased.

Previous research also suggested that use of science writing heuristics (SWH) is facilitator for students' learning (Akkus et al., 2007; Chen et al., 2016; Cronje et al., 2013; Hand et al., 2004). According to Akkus et al. (2007), SWH is a bridge between personal and scientific knowledge. Likewise, Chen et al. (2016) claimed that SWH assists students to engage in argumentation. By writing their initial ideas on SWH, students start engaging argumentation. After students engage in oral argumentation, they write their final ideas on SWH. The difference between their initial and final ideas shows how much students learn in argumentation process. In this study, SWH student template was actively used from the beginning to the end for each week. For example, students wrote their prior knowledge about content in first prompt. Then, they recorded their research questions, hypothesis, materials, and data collection procedure in SWH template. When students conducted their investigation, they were requested to write their observations in SWH templates. They were also requested to use these observations to prepare their individual arguments. After, students wrote their initial arguments, students had time to construct their group arguments. Next, students wrote group arguments in SWH template. When groups presented their group arguments to the rest of the class, they actively used SWH templates in these presentations. If groups did not prepare SWH templates, most probably they could not present their arguments, data, evidence, experiment process in an effective manner. In other words, SWH templates prepared by students were used as source to engage in argumentation and facilitated their engagement in argumentation process. However, although SWH was actively used throughout the process, it is important to note that, students' written arguments included some deficiencies. In general, students successfully reported their research questions, hypothesis, experiment process, recorded observations and prepared individual arguments. However, students tended not to report the diverging ideas in small group discussions. Likewise, students did not report the different ideas revealed in whole class discussions in their SWH templates. Similarly, their reflection was naïve. In the present study, results about written arguments existed in SWH templates were not used as data sources and their results were not shared in result section. The informal examination of SWH written arguments, on the other hand, showed that students did not benefit from SWH student template so much even though they were encouraged to use SWH template in whole process. If students could use SWH better, their written arguments might have been better. Improvement in written arguments could also result in better content knowledge.

Previous research also indicated that the group discussions in ABI process can assist in conceptual understanding (Hand et al., 2004; Heng et al., 2015; Shamuganathan & Karpudewan, 2017). For example, Heng et al. (2015) claimed that group arguments have higher quality than individual arguments because there is a knowledge pool including different ideas in group argument and students integrate their cognitive strengths in this pool. Moreover, students adjust their ideas using peer feedbacks. As a result, their content knowledge improves. Similarly, Hand et al. (2004) asserted that students learn more when they engage in discussion because they learn by listening others' ideas, sharing ideas, commenting on ideas and sharing cognitive load in group discussion. In this ABI research, participants mainly engaged in discussions. Discussions included pre-activity discussions, small group discussions, and whole class discussions. Pre-activity discussions were held at the beginning of the weeks as warm up activities and to reveal participants' prior knowledge in a given topic. Small group discussions were carried out while students were forming their arguments based on their investigations before students present their arguments in whole class discussion. Whole class discussions were done when groups presented their arguments to rest of the class and critically evaluated other groups' arguments. It is highly possible that such activities supported students' conceptual understandings. For example, a concept cartoon was used in second week pre-activity. In this cartoon, a snowman was shown to students and whether snowman without clothes or snowman wearing coat melts earlier in a hot day was asked to students. Most of the students thought that snowman wearing coat melts earlier than snowman without clothes. This shows that majority of students had naïve ideas about heat insulation at the beginning. After that, students engaged in discussions in this pre-activity. The ones who knew heat insulation correctly persuaded others in this discussion. They claimed that snowman wearing coat does not melt early because coat on snowman blocks heat transfer. Most of the students found this argument acceptable at the end of pre-activity discussions and most students reached scientific explanation of the phenomenon about melting of snowman. The increase in the number of students who thought that snowman without coat melts earlier shows that students' content knowledge about heat insulation increased in this pre-activity discussion.

Another factor increasing students' content knowledge might be related with applying ABI phases strictly. For example; Cronje et al. (2013) did not apply these phases strictly and experiment group did not outperform control group regarding content knowledge. Similarly, Rudd et al. (2001) used only SWH student template as laboratory report format and did not focus on other aspects of ABI for experiment group. Control group got traditional instruction and completed standard laboratory report. Groups did not differ in terms of content knowledge (Rudd et al., 2001). These findings showed that if ABI approach is used with its all phases, its efficacy becomes top level; however, ignoring some aspects of ABI (e.g. using only SWH template) or not focusing its phases might decrease the effectiveness of ABI. In other words, all phases of ABI appear to contribute to students' content knowledge and none of the phases should be ignored. Consistent with this idea, all phases of ABI were implemented in the current study. For example, lesson plans were prepared based on these phases for all weeks. Each phase included specific activities. Therefore, both implementing teacher and participants engaged in all activities throughout the study. For example; Benjamin Franklin experiment about lightning was discussed with students for the phase "eliciting students' prior knowledge" in week 5 (i.e., electrical conductivity). Then, materials were provided to students and students designed their experiments in phase "prior activities". Next, students conducted their experiments in phase "laboratory experiments". After that students engaged in negotiation phases 1-2-3-4. In these phases, students constructed their individual arguments and group arguments. They engaged in whole class discussions and compared their findings with authorities like teacher and textbooks. Moreover, students prepared their SWH templates throughout the process. To sum up, all phases of ABI were considered in this study and this consideration might have contributed to the improvement of students' content knowledge.

Lastly, according to the relevant literature, students' identification of their own research questions and their follow of their own data collection procedures can increase the effectiveness of ABI approach (Cronje et al., 2013; Hand et al., 2004). In fact, Hand et al. (2004) claimed that when students formulate their own research questions, their interest increases and they become more likely to complete their research. Similarly, Cronje et al. (2013) reported that pre-determined hypothesis and data collection decrease the effect of ABI because students think that there is one right answer which is the result of testing given hypothesis and following given data collection process. Moreover, students start to think that this is not their own research and their interest decrease towards conducting investigation. As a result, they cannot learn the topic. In this study, students were asked to prepare their own research question and data collection process; however, all lessons were prepared based on curricular objectives. When students prepared research questions extending curricular objectives, they had to modify their research question. Likewise, materials used in this study were easily accessible school materials. Sometimes, materials were not sufficient to conduct students' own investigations and therefore, students had to change their research questions and data collection process. Although, such mandatory changes might have negative effect on students' interest interfering with better content knowledge acquisition, the overall effect of the ABI on students' content knowledge was positive as indicated by a significant increase in mean scores from pre-test to post-test with large effect sizes.

In conclusion, in the current study, there was a significant improvement students' content knowledge. This finding can be attributed to the focus of the ABI on the process of science, the implemented activities' consistency with conceptual change approach (e.g., discrepant events like heat transfer experiment), active use of SWH student template, emphasis on discussions including small group, pre-activity, and whole class discussions, and application of all phases of the ABI strictly. In addition, the effect size was found to be larger for the second topic (i.e. electricity) compared to the first topic (i.e. matter and heat). This finding might suggest that students got familiarity with the ABI and become more adapted to its implementation in the second topic, leading to a better improvement in their content knowledge. Thus, this explanation may suggest that as the ABI becomes a regular part of instruction and students gain more experience, the effectiveness of the ABI on the content knowledge may become more striking.

### 5.1.2. Argument Based Inquiry and Epistemological Beliefs.

According to Conley et al. (2004), epistemological beliefs are the belief systems including source, certainty, development and justification dimensions. Source refers to where scientific knowledge comes from, certainty concerns whether scientific knowledge is certain, development is related with whether scientific knowledge changes, and justification emphasizes role of experiment, testing and evidence in science. In the present study, the EBQ was used to assess students' epistemological beliefs. Findings of this study suggested that there was no significant change in participants' epistemological beliefs in ABI treatment. There may be different explanations for this result: in this study, participants' mean scores for epistemological beliefs were around 4.0 out of 5.0 in pre-tests. Therefore, it was difficult to obtain a significant improvement in participants' epistemological beliefs throughout the study.

There can be some other reasons explaining why participants' epistemological beliefs did not change. For example, lack of explicit instruction about epistemological beliefs might have caused a non-significant change in students' epistemological beliefs. McDonald (2010) reported that explicit NOS (e.g., part of epistemology) instruction is needed to improve participants' NOS views and argumentation treatment without explicit NOS instruction is not sufficient alone to increase participants' NOS views because participants might not link their performance in argumentation and related epistemological beliefs. On the other hand, explicit instruction might assist them to link their scientific practice and epistemological beliefs. By this way, they might increase their epistemological beliefs (McDonald & McRobbie, 2012; McDonald, 2010). However, current study did not include explicit instruction about epistemological beliefs and it is possible that participants did not develop their epistemological beliefs because of this reason.

Likewise, previous studies reported time is important factor affecting quality of argumentation studies (e.g., Chen et al., 2016). If durations of the treatments are longer students benefit more from argumentation studies. However, current study lasted six weeks and six weeks may not be sufficient to improve participants' epistemological beliefs.

According to current study, pre-test results showed that participants of the study held sophisticated epistemological beliefs because students had high scores in epistemological belief items. However, this situation might not reflect the reality. According to Cavagnetto et al. (2010), science classes mainly focus on product of science that is content knowledge and they ignore science practice. This focus on product of science causes students having naïve understandings about science. They see science as collection of facts which is consistent with positivist (e.g. less sophisticated beliefs) epistemology. Therefore, their epistemic commitments might be problematic and this might decrease quality of arguments. Actually, students' proximity to positivist epistemology can be also true for this study although their responses to epistemological belief questionnaire supported that students had sophisticated epistemological beliefs. First of all, participating teacher and students focused on understanding content knowledge prior to the study. Although teacher said they usually discuss on topics, they hardly conducted investigations. Therefore, both teacher and her students were now knowledgeable about scientific practice and most probably, they did not held sophisticated epistemological beliefs prior to the study. Likewise, observational data obtained from students' engagement in argumentation process supported the view that participating students might have naïve epistemology (e.g. positivist epistemology, less sophisticated beliefs). Kolsto and Ratcliffe (2007) claimed that students usually use data coming from expert opinion. Students think that expert opinion is perfect and objective and they do not criticize the expert opinion. In this study, students were more active in the weeks where first hand data were not collected. Students actively used expert opinions existing in evidence cards. While they proposed their arguments, students did not criticize the evidence cards and this is one of the evidence that students had naïve epistemologies. Furthermore, less active class engaged in argumentation less in the weeks that students conducted experiments comparing with the more active class. According to Wallace (2004), students who use only first-hand data or second-hand data held naïve epistemological beliefs. Wallace claimed that students having developed epistemological beliefs use both first hand data and second hand data and they link these data sources in argumentation. Therefore, less active class's focus on second hand data ignoring first hand data during argumentation might supported the idea that participants had naïve epistemologies.

Moreover, McDonald and McRobbie (2012) claimed that transition of epistemological beliefs for a person is related with developmental process. Accordingly, young people think that knowledge is certain, therefore, they are absolutist that is consistent with positivist epistemology (e.g. less sophisticated beliefs). When these people reach to adulthood, absolutist epistemology replaces with evaluationist epistemology that evaluates evidence and prefers more justified models which is consistent with constructivist epistemology (sophisticated beliefs). Therefore, participants of this study ($6^{th}$ grade students in 12 years old) might have positivist epistemology prior to the study.

If participants had positivist epistemology, why did they show themselves in constructivist epistemology? The first answer for this question can be related with students' perception. It is possible that students could not link work of scientists and their work in class. While their epistemology is constructivist regarding scientists' works, they might have positivist epistemology about themselves working in class (Sandoval & Millwood, 2008; Tucel, 2016; Tucel, Cakiroglu, & Öztekin; 2015). Therefore, they might represent constructivist epistemology thinking scientists' work instead of thinking themselves. Second answer can be more philosophical. Students' markings to epistemological beliefs items might not show that these students had constructivist beliefs. On the other hand, their markings might show that these students only agreed the ideas supporting constructivist epistemology, but their personal epistemologies might be different from this constructivist epistemology. Third answer to this question might be about multiple epistemological beliefs that students held. In one study conducted by Liu and Roehrig (2019) showed that participants might have two different epistemological beliefs in same time which are general epistemologies and topic-specific epistemologies. Liu and Roehring (2019) found that participants' general epistemology was constructivist because participants believed that scientific knowledge is uncertain. On the other hand, participants thought that the findings about climate change were "truth". This showed that participants' topic-specific epistemological beliefs were positivist. Similar result can be true for current study too. For example, participating students might have general epistemology consistent with constructivist epistemology. Therefore, they might rank the items supporting the constructivist epistemology in pre-test and post-test. On the other hand, participants might hold positivist epistemology in specific science topics (e.g., electricity) that they engaged in argumentation and they acted in class consistent with positivist epistemology. For instance, when participants reached different conclusions in salty water experiment, they thought whether their experiments or other groups' experiments were wrong, but students did not think the conditions leading different results in their experiments. The idea of wrong or true result might show that students had positivist epistemology in electricity topic. In conclusion, their general and topic-specific epistemologies might be different from each other.

If participants had positivist epistemologies as explained above prior to this study, this might also cause no change on their epistemological beliefs. Accordingly, Walker and Sampson (2013) sophisticated epistemological beliefs are necessary to construct qualified arguments because the ones having sophisticated epistemological beliefs evaluate different evidence and they are open to different ideas and evidence. In other word, the ones having sophisticated epistemological beliefs are ready to engage in argumentation. On the other hand, positivists believe that there is one correct answer and there is no need to discuss more after reaching correct answer. Therefore, if participants of current study had positivist epistemologies at the beginning of the study, this might decrease their engagement in argumentation, and decreased engagement in argumentation might inhibit development of participants' epistemological beliefs.

In conclusion, findings of the study suggested that participants of the study had constructivist epistemological beliefs before and after the study based on quantitative analysis and participants' epistemological beliefs did not change in this study. There can be different reasons why students' epistemological beliefs did not improve in this study. It is possible that high scores prior to the study, lack of explicit epistemological belief instruction in ABI treatment, and short duration of ABI treatment might cause no change on students' epistemological beliefs in this study. Although quantitative results showed participants had constructivist epistemology (e.g. sophisticated beliefs) participants might not have constructivist epistemological beliefs. On the other hand, they might have positivist epistemology (e.g. less sophisticated beliefs). Participants' traditional instruction focusing on product of science prior to the study, students' approach to expert opinion (i.e., not criticizing evidence cards), less active class's less engagement in argumentation in the weeks they conducted experiment, participants' age and its corresponding epistemological beliefs aligned with developmental perspective supported the view that participants might have positivist epistemologies. Students' inability to link their work and scientists' works, the difference between their actual epistemologies and their approval of constructivist epistemology, and their multiple epistemologies including general and topic-specific epistemologies might cause students show themselves having constructivist epistemologies although they might have positivist epistemology. Following part continues with discussing findings of current study and previous research about science process skills.

#### 5.1.3. Argument Based Inquiry and Science Process Skills.

One of the aims of current study was eliciting the change of science process skills in the ABI treatment. Padilla (1990) claimed students' science process skills increases when they engaged in science practice. According to Sanderson and Kratochvil (1971), there are two different kinds of science process skills which are the result of classification of science's intellectual tool: basic skills (e.g., observe) and integrated skills (e.g., defining operationally). While basic skills are taught in primary years of schooling, integrated skills are taught in the following years. Consistent with the participants' grade level (i.e. $6^{th}$ grade), integrated skills represented students' science process skills and test of integrated process skills developed by Burn et al. (1985) was used to assess participants' science process skills.

Due to fact that reliability of test scores was too low for components of integrated process skills (e.g., identifying variables), the test was used as unidimensional. The same issue was also reported in previous research and the scores from the overall instrument were utilized in data analysis (e.g., Gök, 2014). In the present study, findings suggested that participants' science process skills as a unidimensional construct did not change significantly in this treatment. There can be different reasons why participants' process skills did not improve in this treatment. These possible factors may be short duration of the treatment and lack of explicit emphasis on science process skills throughout the study.

Regarding the duration of the treatment, although the treatment lasted six weeks, students conducted experiments which involved the extensive use of science process skills only in four weeks. Thus, this situation might have led to a non-significant change in students' science process skills. Similarly, Yıldırım (2012) reported that short period of guided inquiry treatment caused no change on students' science process skills. If this study had lasted longer and students conducted experiments in all weeks, it would be possible that there would be an improvement in students' science process skills. At least, theoretical consistency between ABI and science process skills promises this improvement.

Why participants' science process skills did not improve can also be related with lack of emphasis on science process skills during treatment. Although students used science process skills in this treatment especially in the weeks they conducted experiment, science process skills were not emphasized in these experiments. For example, in their intervention, Durmaz and Mutlu (2017) emphasized science process skills in experiment group, and experiment groups' science process skills developed more comparing with control group in which science process skills were not emphasized. In another study, Akben (2015) asked participants which science process skills they used when they designed and conduct their experiment in an inquiry study. By this way, participants became alerted about their science process skills in their investigation and they could learn science process skills. Moreover, participants discussed about science process skills with their peers and peer feedbacks probably improved their process skills further. If participants were asked to address which science process skills they used in their experiments in this study, it is possible that there would be an improvement their science process skills. But, in the current study, neither science process skills were emphasized nor there was a discussion on science process skills used and these two situations might have caused no change in participants' science process skills.

To sum up participants' science process skills did not improve in this study and possible reasons for no change on science process skills can be short period of treatment time and lack of emphasis on science process skills during treatment.

Until now, how participants' content knowledge, epistemological beliefs and science process skills which are components of scientific literacy were affected by argument-based inquiry treatment was discussed. The next part of discussion is about argumentation schemes students used in this study.

### 5.2. Nature of Argumentation Schemes in Middle School Level

Walton's (1996) argumentation schemes and Duschl's (2007) explanations about argumentation schemes were used to reveal the nature of 6$^{th}$ grade students' arguments. According to Walton et al. (2008), argumentation schemes are forms of arguments showing structures of common types of arguments people use in everyday discussions. Argumentation schemes are presumptive in nature and they are used to support main arguments. This support can assist to change direction of discussion. In other word, the weight that argumentation schemes have contributes on other arguments that schemes are related.

In this study, five assertions were proposed about 6$^{th}$ grade students' argumentation schemes. Three of these assertions were derived from similarities of two classes observed and two of assertions were derived from differences of two classes. In this part, these five assertions are discussed.

Firstly, this study showed that number of types of argumentation schemes students used increased over time. Secondly, students preferred to use some argumentation schemes like argument from sign more than others. Third, the data sources that students used affected which argumentation schemes students used. Fourth, the more active class used the more types of argumentation schemes. Fifth, the more active class used argument from position to know most, but the less active class used argument from expert opinion most.

More specifically, this study, firstly, showed that there was an increase the total number of type of argumentation schemes used in both class 1 and class 2 over time. Accordingly, total types of number of argumentation schemes increased from 5 to 8 for class 1 and total types of number of argumentation schemes increased from 7 to 12 for class 2. On the other hand, there were fluctuations for total number of argumentation schemes used. For example, the number of argumentation schemes used in class 1 students was the highest in the week 3, week 4 and week 5, but fewest number of argumentation schemes were used in week 1, week 2, and week 6. These findings suggest that although total number of argumentation schemes used fluctuated, total types of number of argumentation schemes increased over time in ABI treatment. This finding may imply that as students get familiarity with the argumentation over time, they start to use different types of argumentation schemes. According to Konstantinidou and Macagno (2013) the use of argumentation schemes as a way of analysis makes participants' reasonings concrete. In other word, argumentation schemes that participants use show their reasoning. Therefore, increase of total types of argumentation schemes supports the idea that participants do more reasoning over time. Similarly, in another study using argumentation schemes, Macagno et al. (2015) reported that students considered values more on their arguments and they offered more solutions to the problems over time. They focused on the reasons underlying in an argument rather than argument itself in their critics over time. Therefore, it can be said that participants use more types of argumentation schemes over time in ABI reflecting a development in their reasoning. On the other hand, Duschl (2007) claimed that nine of argumentation schemes which are sign, commitment, position to know, expert opinion, evidence to hypothesis, corelation to cause, cause to effect, consequences, and analogy are appropriate for middle school students, but this study showed that students can use more than these nine argumentation schemes. In this study, middle school students used 17 different types of argumentation schemes (e.g. argument from established rule). This result is evidence that middle school students can use more argumentation schemes than they are expected to use if students have enough opportunity to engage in argumentation as in argument-based inquiry treatment. Moreover, use of various types of argumentation schemes shows that middle school students can do more presumptive reasoning when they engage in ABI treatment. Therefore, it can be said that ABI treatment is an efficient way of improving participants' reasoning in middle school level.

Second, this study showed that participants tended to use some argumentation schemes more than others. More specifically, both classes used argument from expert opinion, argument from sign, argument from position to know, argument from comparison and argument from evidence to hypothesis frequently and they used other argumentation schemes less. Students mainly used argument from expert opinion when they had second hand data like evidence cards. They used argument from sign when they looked for explanation of observed phenomena as explained by Duschl (2007): When students used argument from sign, their explanation sometimes included wrong content knowledge, but students were aware that their explanations were not certain, they included possibility. Classes also used argument from position to know to get further information and to show inconsistencies of ideas. Argument from comparison was mainly used to compare concrete materials consistent with its definition (Walton, 1996). Lastly, argument from evidence to hypothesis was used when students reported their experiments and hypothesis. At this point, it is important to note that in the present study, argumentation schemes were obtained only from whole class discussions not from experiments or small group discussions. However, argumentation schemes derived from experiments can be different from schemes obtained from whole class discussions. For example; Ozdem et al. (2013) examined pre-service teachers' argumentation schemes in an ABI treatment and researchers reported that participants used argumentation schemes to construct their arguments in experimentation and they used argumentation schemes to criticize others in whole class discussions. Ozdem et al. (2013) also reported that argument from sign, argument from evidence to hypothesis and argument from corelation to cause are scientific argumentation schemes because argument from sign is related with inference, observation, measurement and graphs; argument from evidence to hypothesis is used when participants make verification and falsification; argument from corelation to cause is used when participants observe relationship between variables. Argument from evidence to hypothesis and argument from sign which are scientific schemes were also commonly used in current study too, but participants of this study did not use argument from corelation to cause so much. In addition to these schemes, participants of this study used argument from expert opinion, argument from comparison and argument from position to know frequently. In addition to argumentation schemes Ozdem et al. (2013) reported, these three schemes frequently used in current study (e.g. argument from expert opinion) can also be scientific argumentation schemes. For example, students in this study frequently used argument from expert opinion. Wallace (2004) claimed that students should integrate first hand data derived from their experiments and second hand data obtained from experts (e.g., textbook) to construct their explanations. Therefore, students should actively use second hand data by proposing argument from expert opinion in discussion. In this point, using argument from expert opinion can be seen as an example of scientific argumentation scheme. Likewise, argument from comparison can be a scientific scheme. While argument from analogy is used to compare abstract things, argument from comparison is used to compare concrete things (Walton, 1996). Argument from comparison was actively used in current study. Due to fact that participants were expected to be in concrete operational stage, they usually reported macro level things and their relationship in their arguments. By comparing these different observable entities, participants used argument from comparison. Likewise, participants actively used argument from comparison when they reported the differences between their experiment group and control group in investigations; therefore, argument from comparison can be a scientific argumentation scheme too. Lastly, participants of this study frequently used argument from position to know in this study; however, pre-service teachers did not use this scheme so much in previous research (Ozdem et al., 2013). As other four schemes, argument from position to know can also be scientific argumentation schemes. When participants thought they did not have enough information or they observed something that was inconsistent with their thinking, they preferred to ask questions to presenter groups. By this way, they used argument from position to know. Lederman and Lederman (2012) claimed that scientific inquiry starts with asking questions at the beginning of the scientific investigation and scientists try to answer these questions in their investigations. Similarly, argument from position to know stimulated the argumentation process in class and students looked for answers to these questions. Due to fact that argument from position to know was used to start scientific argumentation in class, this scheme can also be seen as scientific argumentation scheme. Although these five argumentation schemes revealed more than other twenty argumentation schemes, we do not claim that middle school students always use these argumentation schemes and ignore others. For example; if students do not know the content, most probably they are going to create their own explanations and the number of use of argument from example and argument from analogy will increase. Likewise, if the topic is too complex and needs to be broken down before understanding it, it is possible that students will use argument from gradualism frequently. Simlarly, if students have misconceptions in a science topic, it is possible that they are going to link inconsistent things and so they will use argument from corelation to cause in general. However, in this study, use of five argumentation schemes explained above was more frequent than others and so this finding was reported as argumentation schemes assertion 2. This assertion is tentative as all assertions and can change when new evidence emerges.

Third, this study showed when type of data sources (e.g., first hand) change, type of argumentation schemes that students used also changes. Accordingly, when participants conducted their own investigation, they produced first hand data and both classes used argument from evidence to hypothesis by reporting their hypothesis and variables. Students did not use argument from expert opinion in these weeks. On the other hand, participants stopped using argument from evidence to hypothesis when they did not conduct experiment and evidence cards were provided them. In these weeks, participants used evidence cards as second hand data and frequently used argument from expert opinion. In other word, the type of data participants used shaped their use of argumentation schemes. Similarly, Ozdem et al. (2013) reported that participants' use of argumentation schemes changed depending on the tasks. While students used argument from sign most in some tasks, they used argument from corelation to cause most in some others. Likewise, Kind et al. (2011) reported that students do not have certain orientation. They develop their orientation spontaneously. If they conduct their experiments, their orientations is based on experimenting and they focus on data collection; on the other hand, if they do not conduct experiment, they focus on product of science (e.g., already available data). Focusing on one type of scheme depending on data source can be deficiency because students are expected to use first hand and second hand data together in scientific investigation (Walsh & McGowan, 2017). Using only one type of data can support students' positivist epistemological beliefs. For example, they might think their experiment results are the unique way reaching scientific knowledge if they use only first-hand data. Likewise, students might think that second hand data is objective and perfect knowledge and it should not be criticized if they only use second hand data (Wallace, 2004). Therefore, just focusing on argument from evidence to hypothesis in using first hand data or just focusing on argument from expert opinion in using second hand data is problematic and does not align with constructivist epistemology. One of the reason of students' focus on only one type of argumentation schemes depending on type of data source (e.g., first hand) can be design of current research. In this study, additional evidence cards were not given to students when they conducted their experiments. Students were supposed to use textbook or teacher's ideas to produce argument from expert opinion, but they did not in general. They focused on their experiments and used argument from evidence to hypothesis. If evidence cards had been provided to students in these weeks, they might use argument from expert opinion frequently and they might integrate first hand and second hand data. The same also can be true for the weeks that students used second hand data. Accordingly, evidence cards were given to students in those weeks consistent with the curricular content and students mainly use argument from expert opinion in these weeks; however, students had no opportunity to conduct their own investigations, so it is not surprising that students did not use argument from evidence to hypothesis frequently in these weeks.

Findings also suggested that more active class used more type of argumentation schemes than less active class. Actually, two classes used similar number of types of argumentation schemes in first three weeks of the study. However, more active class used 12 different argumentation schemes, but less active class used 8 different argumentation schemes in last weeks of the study. Sampson and Clark (2011) explained that characteristics of discourse affect group performance. Accordingly, the group who interacts more than others focuses on content, uses much more ideas. More and diverse ideas lead to more discussion. As a result, groups engaging in argumentation more are likely to improve their reasoning. Similarly, argumentation schemes show proponents' presumptive reasoning and this reasoning might improve through engagement in argumentation. As a result, groups that improved their presumptive reasoning might use different types of argumentation schemes. Likewise, the group who engaged more in argumentation might need to look from multiple perspectives to support their arguments and to refute others' arguments in discussion. Looking from multiple perspectives might also cause participants to use various types of argumentation schemes.

The study also showed that while more active class used argument from position to know most, less active class used argument from expert opinion most. This finding might explain why more active class engaged in argumentation more and less active class did not engage argumentation so much. According to Lederman and Lederman (2012), asking question is very important in science and scientific inquiry starts with questioning. After proposing questions, scientists investigate the answers of these questions. In this study, more active class started scientific inquiry like scientists by asking their questions shown in their argument from position to know. By this way, class discussed alternative ideas and other findings they did not consider before. These questions and corresponding answers sustained discussions. By this way, Class 2 became more active class in this study. By asking questions frequently, class 2 engaged in argumentation process like scientists. On the other hand, less active class (i.e. class 1) used argument from expert opinion most instead of argument from position to know in this study. Argument from expert opinion was mainly used in this study when participants used second hand data (e.g. evidence cards). Frequent use of the second hand data suggests that owners of these arguments are interested in product of science and they ignore process of science. In other words, they deal with only one aspect of science. However, science is not limited to its product (i.e. content knowledge, knowledge in evidence cards, second hand data). People should also consider how science is done, so they should consider process of science too. If participants solely focus on content knowledge, they might think that there is no need to discuss because everything is certain and represented in content knowledge that is truth (Kind et al., 2011). This might be the reason why Class 1 became the less active class. According to another point of view, Kim and Song (2006) claimed when students do not know the content, they appeal to authoritarian evidence to hide their ignorance. If this is true for less active class, this shows that this class did not reach to threshold level content knowledge to engage in argumentation process. Previous research showed that middle school students may not reach to threshold level content knowledge and therefore, they could not improve rational reasoning (Dawson & Venville, 2013). If these students have lack of content knowledge and rational reasoning, it is expectable for them not to engage in argumentation process. Furthermore, appealing to expert opinion and using mostly second hand data may decrease their interest to engagement in argumentation. Accordingly, Kind et al. (2011) claimed that if students use the data they produced (first hand), they own this data, propose more hypotheses and interpret the evidence further. As a result, they actively engage argumentation process. On the other hand, class 1 may not motivate themselves because they appealed to second hand data which were not belong to them, so they might not own these data, need to support it, and engage in argumentation actively.

In this part, assertions about argumentation schemes that middle school students used were interpreted by considering previous research on argumentation. The next part addresses the discussion for last research question which is about middle school students' nature of engagement in argumentation process.

## 5.3 Nature of Students Engagement in Argumentation Process

Students' engagement in argumentation process was revealed using codes obtained from the Sampson and Clark's (2011) study. Accordingly four codes which are expositional comments, oppositional comments, information seeking, and co-construction of knowledge were used to reveal assertions about students' engagement in argumentation process. Expositional comments were used when participants proposed, clarified and justified their own ideas. Oppositional comments were used when participants challenged others' ideas. Information seeking is used when there is no enough knowledge and participants asked further clarification. Co-construction of knowledge was used when participants summarized, revised, supported or added to the others' ideas (Sampson & Clark, 2011).

Similarities and differences between two classes revealed eight assertions. Seven of these assertions emerged from commonalities of two classes and one assertion was derived from differences between two classes regarding students' engagement in argumentation process. Three of the seven assertions derived from commonalities are about nature of the activities used in this study. Accordingly, assertions about nature of activity showed that "Students engage in argumentation at most when evidence cards are provided to them.", "Students engage in argumentation less in a following week when they do similar activities.", and "Use of ideal conditions might inhibit students' engagement in argumentation.". On the other hand, four of the assertions derived from commonalities are related with the components about students' engagement in argumentation process. These four assertions showed that "When students engaged in argumentation, they mostly use expositional comment.", "If evidence cards are provided to students and students have prior knowledge about the topic, they mostly tended to use oppositional comments in argumentation process.", "Students use information seeking component more in the weeks they conduct experiment comparing with the weeks they do not conduct experiment. Moreover, use of information seeking becomes top level when students conduct experiment at the beginning of the units.", and "Students usually do not use co-construction of knowledge in argumentation." Finally, one assertion about students' engagement in argumentation process emerged from the differences between two classes and this assertion claimed that “While more active class tend to engage in argumentation process more in the weeks when they conduct experiments, less active class tend to engage in argumentation process more in the weeks that students use evidence cards to construct arguments.”

In this part, firstly, assertions about nature of activities are discussed. Secondly, assertions about the components related with students’ engagement in argumentation process are discussed. Then, assertion derived from differences of two classes is discussed.

In this study, first assertion about nature of activity is “Students engage in argumentation at most when evidence cards are provided to them.” Previous research also showed that students use already available data, and they rarely use their own data (Kolsto & Ratcliffe, 2007; Walsh & McGowan, 2017). The possible reason for this situation regarding why students tend to use evidence cards rather than their own data can be related with the distinction between data and evidence. Accordingly, evidence is the transformed data that is directly used to construct argument. On the other hand, raw data cannot be used directly to support a claim; it needs to be transformed to evidence. Transforming raw data into evidence is more difficult than using already available evidence cards, so students might usually prefer to use evidence cards and so their engagement in argumentation was more in the weeks when they used evidence cards (i.e. week 3, week 4). Moreover, Kolsto and Ratcliffe (2007) claimed that people notice the expert authority more than data and this situation become obvious when time, content knowledge and skills are limited. In this study, whole class discussions lasted 40 minutes. If the discussion time was more, students might have focused on their own data more because they would have more time to think on their data, and so they might have used their own data more to construct arguments. Likewise, lack of content knowledge might have resulted in students’ reliance to evidence cards. This was especially case in week 3, thermal insulating products. In this week, students had no prior knowledge and they had lack of knowledge. Their arguments mostly depended on evidence cards in this week. Likewise, Kim and Song (2006) claimed when students do not know the content, they use authoritarian evidence to hide their ignorance. Similarly, argumentation was new for the participants of the study and the study lasted 6 weeks. It is possible that this duration was not sufficient to increase students' argumentation skills such as connecting data and claim. In this study, the relation between evidence and claims were obvious and it did not require so much cognitive demand. On the other hand, students' raw data were unclear and connecting them to claims required high cognitive demand and developed argumentation skills. If students did not have developed argumentation skills, they might tend to connect evidence to claims instead of connecting raw data to claims because of their limited argumentation skills. Other two reasons why students were more active when using evidence cards can be related with reasoning and the social costs those students pay (Zemplen, 2011). According to Zemplen (2011), students do not make reasoning in class so much and this lack of reasoning might inhibit their engagement in argumentation. It is possible that if students did not make reasoning in class, they directly used already available evidence cards to support their claims because using already available evidence cards require less reasoning comparing with using raw data to support claim. Likewise, Zemplen (2011) claimed that proposing own idea is riskier than using expert authority. Class is a social environment and if students' own ideas are wrong, students might think that they will be blamed by peers and they will pay the cost of their wrong ideas. In order to eliminate this possible risk, students might not propose their ideas so much, and they preferred to use evidence cards.

Second assertion about nature of activity released in this study is "Students engage in argumentation less in a following week when they do similar activities." This study showed that task similarity between different weeks decrease students' engagement in argumentation process. In this study, week 1 focused on heat conductivity and week 2 focused on heat insulation and both topics were about heat transfer. In these weeks, students might think they repeated the same activities and they might lose their excitement. This situation can be explained by unsurprising stimulus effect (Chin & Osbourne, 2010). Chin and Osbourne (2010) advocated that a stimulus is given in argumentation at the beginning. This stimulus is a problem and source of data. If this stimulus is not surprising as in the week 2 because it was similar to week 1 activity, people think that this stimulus is unnecessary, therefore, people do not ask questions about stimulus and they do not elaborate their arguments. On the other hand, if stimulus is surprising, students experience cognitive conflict and they start to think about problem (i.e. stimulus). There is a cognitive dissonance between existing ideas and new situation. Student asks question to herself/himself and looks for the answers. Then, student adjusts existing ideas consistent with the data emerged from stimulus. This is the monological argumentation that student does herself/himself. After that these monological arguments are shared with peers a dialogical argumentation starts (Chin & Osbourne, 2010). However, surprising stimulus sometimes does not trigger to argumentation, if students reject or ignore anomalous data (i.e. surprising stimulus) (Chin & Osbourne, 2010). To sum up, task similarity between week 1 and week 2 acted as unsurprising stimulus and this situation might decrease students' engagement in argumentation process. Therefore, selecting different activities which were highly different from week 1 activity and surprising for week 2 could have been better to engage students to argumentation more. For example, students could have designed and prepared a thermos in week 2 which is about heat insulation. If students had designed and prepared thermos in this week, most probably making thermos would have been different from first week activity and got students' attention more as surprising stimulus. Therefore, students might have engaged in argumentation more if they had designed a thermos.

Third assertion about nature of activities found in this study is "Use of ideal conditions might inhibit students' engagement in argumentation". Laboratory experiment was replaced with simulation activity in last week because experiment results were faulty in pilot study. Students reached correct results by using simulation activity; however, they did not engage argumentation so much because conditions became idealized in simulation which means there was no error. Idealized conditions caused similar results, and students' explanations were similar too. Hence, students did not discuss further in this week. According to Jimenez-Aleixandre (2008), curricular programs should include activities including dilemmas, and students should reach different results in tasks and this situation should be explained by different ideas. However, use of simulation activity eliminated dilemmas, different results and different interpretations in this study.

Moreover, what researcher sees as outcome in this study might cause participants' less engagement in argumentation in last week of main study. In pilot study, researcher thought outcome of the study is reaching correct content knowledge for the students. Students could not reach correct knowledge in pilot study, but they actively engaged in argumentation. Therefore, researcher decided to replace experiment with simulation activity to get students to reach correct knowledge in main study. As it was planned, students reached correct knowledge in main study; however, students did not engage in argumentation so much because simulation activity did not lend itself any error, mistake and diverging ideas. The outcome of main study was students' engagement in argumentation and this outcome failed in some extent when simulation activity was replaced with experiment for the last week of the study. In this point, Sampson and Clark (2011) claimed that what researchers perceive as outcome is highly related with students' engagement in argumentation. For example, use of oppositional comments is the characteristic of scientific explanation if researcher's outcome is to make students produce high quality of arguments. On the other hand, if researchers see scientific inquiry that students do in argumentation as outcome, results can be different regarding students' engagement in argumentation (Sampson & Clark, 2011).

This study suggested that nature of activity (e.g. use of evidence cards, unsurprising stimulus, and idealized conditions) is one important factor affecting students' engagement in argumentation process: Because of nature of activity, a trend regarding students' engagement in argumentation process might not be observed. For example, unsurprising stimulus decreased students' engagement in argumentation and this was the second week; hence an improvement in the students' engagement was not observed in this week for both classes. Likewise, evidence cards were used in week 3 and week 4. Students' engagement became top in these weeks when they used already available evidence and so their engagement in argumentation did not further improved in last two weeks (i.e. week 5 and week 6). Similarly, students met with idealized conditions when they used simulation program in last week of the study, and idealized conditions decreased their engagement in argumentation process. On the other hand, it was expected that students should have engaged in argumentation most in the last week because students had six weeks of experience regarding argumentation in this week. To sum up, nature of activity affected the change of students' engagement in argumentation process throughout the study. This situation interfered with determining whether students' engagement in argumentation develop over time in this study. On the other hand, previous ABI studies examining students' engagement in argumentation process revealed that participants engaged in argumentation more over time (Chen et al., 2016; Walker & Sampson, 2013). For example, Chen et al. (2016) reported that fifth grade students used more aspects of social negotiation over time. While students used information seeking and elaboration at the beginning of the study, they started to use challenge, reject, support, and defend aspects in the following weeks. While students focused on constructing arguments in first weeks, they focused on both construction and criticizing arguments in their engagement to argumentation (Chen et al., 2016). In their study, researchers' study lasted eight weeks for each unit which were living things and their environments and human body systems. Researchers did not report any problem decreasing students' engagement in argumentation process unlike current study. First activity was about the factors affecting germination of seeds and the activity lasted six weeks. Students asked their own research questions and followed their investigation to answer research question. Second activity lasted two weeks was about living organisms' effect on their environment. Students made ecological system map and discussed, then they examined ecological balance considering a deer activity. Third activity was about working mechanism of digestive system lasting four weeks. Students' digestive system activity includes demonstration of how digestive system works. Last activity was about working mechanism of respiratory system and lasted four weeks. Students built a model to understand how respiratory system works. Some differences between Chen et al.'s (2016) study and current research might cause the different results in terms of students' engagement in argumentation. Firstly, Chen et al.'s (2016) study lasted 16 weeks, but current study lasted 6 weeks. Chen et al.'s (2016) study's participants got more experience to ABI and therefore it is expected that these students' engagement in argumentation increased over 16 weeks. On the other hand, participants of the current study had six weeks of experience that is less experience relatively and so their engagement to argumentation process may not improve extensively. Moreover, while idealized conditions and unsurprising stimulus interfered students' engagement in argumentation process in this study, such problems were not reported in Chen et al.'s (2016) study. However, comparison of two studies showed one major difference that is independence of researcher. Two contexts were highly different from each other. In Chen et al.'s (2016) study curriculum had no pressure on researchers. For example, curricular objectives were not reported in Chen et al.'s (2016) study. On the other hand, implementing teacher formed big ideas to cover in Chen et al.'s (2016) study. On the other hand, curricular objectives had to be taught in current study. Lesson plans were prepared based on these objectives which might decrease the efficacy of ABI treatment. Likewise, students were completely free in their investigation as long as their investigation is about big idea in Chen et al.'s (2016) study. On the other hand, students of the current study were not totally free because their research questions were needed to be related with curricular objectives. The research questions not matching with objectives were replaced with new questions and these obligations might decrease students' enthusiasm to engage in argumentation process in current study. Walker and Sampson (2013) reached similar findings with Chen et al. (2016) regarding students' engagement in argumentation. Walker and Sampson (2013) claimed that students focused on generating arguments in first weeks, but they started to evaluate arguments in following weeks. Researchers explained this observation based on students' adaptation to argumentation. Accordingly, students could not adapt to argumentation and they could not criticize arguments in initial weeks (Walker & Sampson, 2013). To sum up, while previous studies showed students' engagement in argumentation developed over time in ABI research, this study did not suggest development regarding students' engagement in argumentation process and it is possible that nature of activities caused this situation.

Similar to assertions about nature of activities, this study has four assertions about the components of students' engagement in argumentation process derived from commonalities of two observed classes. These assertions are discussed in this part.

First assertion about engagement components is about expositional comment. The assertion about expositional comment is “When students engaged in argumentation, they mostly use expositional comment”. This assertion showed that students mainly preferred to use proposing, justifying, and clarifying ideas. Likewise, Cavagnetto et al. (2010) reported that fifth grade students spent more time to constructing arguments and less time to evaluating them and students rarely used rebuttals in their arguments. Similarly, Garcia-Mila and Anderson (2007) claimed that students focus on their own ideas and they do not consider others’ ideas and alternative perspectives so much.

Regarding the place of expositional comments in argumentation process, expositional comments can be seen as starting point of argumentation. For example, groups proposed their group arguments at the beginning of whole class discussions proposing their ideas using expositional comments. By referring these expositional comments, other students used expositional comment, oppositional comment, information seeking, and co-construction of knowledge. Similar to this finding, Kim and Song (2006) reported that after focusing on their ideas, middle school students start argumentation by exchanging their ideas as reporters provide information about their research. However, current research provides more information about the status of expositional comments in argumentation process. Characteristics of the activities used in this study and their relationship with expositional comments provide further information regarding nature of expositional comments in current study. Accordingly, students did not use expositional comments when they lost their excitement (i.e. use of unsurprising stimulus) (see Figure 5.1). On the other hand, students used expositional comments most when they had no prior knowledge and evidence cards were provided to them. However, students did not use other three engagement components when they had no prior knowledge and evidence cards are given them. They just focused on constructing their arguments by using expositional comments (see Figure 5.2). These explanations about expositional comments are two instances that students do not engage in argumentation efficiently, and their engagement does not go beyond expositional comments.

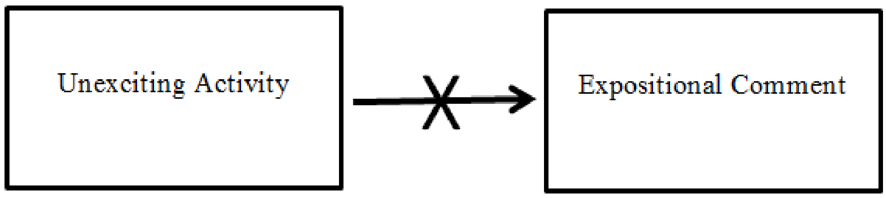



*Figure 5. 1* Non-working argumentation process 1


*Figure 5. 2* Non-working argumentation process 2

Second assertion derived from similarities of two classes regarding engagement components is related with oppositional comments. Use of oppositional comments was seen as characteristics of scientific argumentation in previous research. Accordingly, Sampson and Clark (2011) advocated that critics revealed by use of oppositional comments maintain and deepen discussion. Sampson and Clark (2011) also added that some researchers thought that scientific knowledge is constructed by elimination of some ideas and these ideas are eliminated through critics. Therefore, using oppositional comments can be way of reaching knowledge in middle school level. Current study provides some information regarding when students use oppositional comments frequently. Accordingly, second assertion about students' engagement claimed that "If evidence cards are provided to students and students have prior knowledge about the topic, they mostly tend to use oppositional comments in argumentation process." Both classes used oppositional comments most in week 4 that is about type of energy sources. Students were highly active in pre-discussions in week 4 and this is evidence that students had prior knowledge about content. Moreover, evidence cards were provided in this week and combination of students' prior knowledge and use of evidence cards supported their use of oppositional comments. While students mostly focused on oppositional comments in this week, other aspects of their engagement in argumentation (e.g. information seeking) were used less frequently. Figure 5.3 summarizes when students of current study used oppositional comments most. According to figure 5.3, exciting activity, topics that students have prior knowledge and use of evidence cards form a baseline for students to engage in argumentation. However, oppositional comments does not reveal directly. Before occurrence of oppositional comments, students propose their ideas using expositional comments as starting point and after that oppositional comments are mostly observed during argumentation process. Naylor, Keogh and Downing's (2007) explanations about how students view argumentation also support this assertion. Accordingly, there are two conditions determining students' perception towards argumentation. These conditions are "not knowing the content" and "knowing the content". If students do not know the content, argumentation becomes a tool for them and the aim is reaching knowledge, so students use co-construction of knowledge to reach knowledge. On the other hand, if students know the content, argumentation becomes aim for students, and confrontations dominate the class, therefore students frequently use oppositional comments (Naylor et al., 2007).

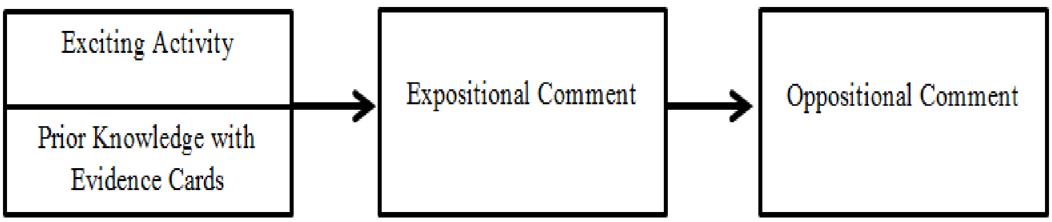



*Figure 5. 3* Schema for when students use oppositional comments most

Third assertion derived from similarities of two classes is about information seeking. This assertion claim that "Students use information seeking component more in the weeks when they conduct experiment compared with weeks when they do not conduct experiment. Moreover, use of information seeking becomes top level when students conduct experiment at the beginning of the unit." Accordingly, information seeking was generally used at the beginning of unit and when students conducted experiment; however, it was rarely used when evidence cards were given. This finding suggested that information seeking is generally used when students did not know so much about topic (e.g. heat and matter) because the units were new for them at the beginning. In order to know further about topics, students frequently asked questions. For example, students asked theoretical questions about the results of experiments. By this way they aimed to improve their understandings. Likewise, information seeking was used to eliminate uncertainties. For example, different groups reached different results when they conducted the same experiments. Therefore, they asked questions about the reasons of why groups reached different results. Such questions were asked after presenter group reported their experiment results as expositional comments. After other groups asked questions by using information seeking, presenter group made explanation again as expositional comment and argumentation continued with uses of engagement components (see figure 5.4). In a similar line, Chin and Osbourne (2010) claimed that driving forces for people to ask question are their deficiencies and conflict. These conflicts cause asking questions, thinking process starts and people propose their arguments as response to questions. These ongoing questions that disputers ask each other consolidate their arguments (Chin & Osbourne, 2010). Students' limited use of information seeking when they use evidence cards also supported the idea that students use information seeking when there are uncertainties and they do not know so much. Accordingly, evidence cards both provided knowledge to students and eliminated uncertainties in this study. Therefore, students did not need to use information seeking so much when evidence cards were provided them (see figure 5.5). Similarly, Chin and Osbourne (2010) claimed that if people think there is no problem, people do not ask question and make explanation.

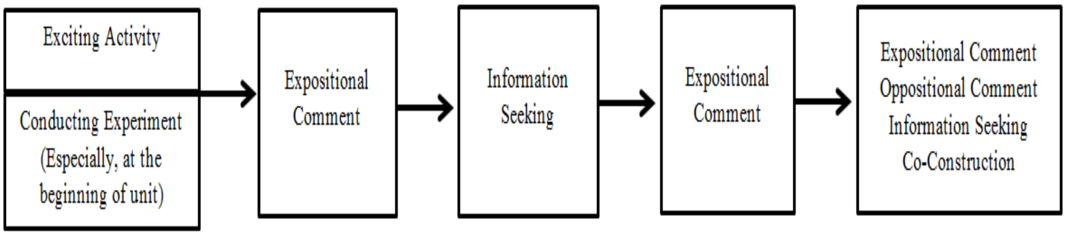



*Figure 5. 4* Schema for when students use information seeking frequently


*Figure 5. 5* Schema for when students do not use information seeking

Last assertion about students' engagement in argumentation process is related with co-construction of knowledge and this assertion claimed that "Students usually do not use co-construction of knowledge in argumentation." Although co-construction of knowledge was not used frequently in this study, there were some instances that students used co-construction of knowledge. Accordingly, participants used co-construction of knowledge in three ways. First, co-construction was used when presenting groups had difficulty. For example, one group had difficulty in designing experiment, and other groups assisted them. By this way, groups could design experiments. Second, co-construction of knowledge was used when groups reached similar results. In these instances, groups supported each other by claiming they reached the same results. By this way, groups increased the reliability of their results. Third, co-construction of knowledge was used when students did not have enough knowledge to explain observed phenomena. For example, paper cup did not conduct electricity and students asked the reason for why paper cup did not conduct electricity. Students did not know the core idea that is resistance, but they tried to understand why paper cup did not conduct electricity during argumentation. By adding their ideas, students linked heat conduction to the electrical conduction and they reached the explanation that the space between particles caused paper cup not to conduct electricity. In this example, students used co-construction, but did not reach correct core idea (i.e. resistance). Therefore, it can be said that use of co-construction of knowledge does not guarantee to reach scientific explanation. Further information regarding the use of co-construction was given in previous research. Accordingly, Naylor et al. (2007) claimed that co-construction of knowledge is as important as confrontation in science. When students do not have enough information, they try to reach knowledge through co-construction (Naylor et al., 2007) as in the case that students tried to explain why paper cup did not conduct electricity. On the other hand; if students know the content, they compete and eliminate ideas by use of oppositional comments (Naylor et al., 2007). This study showed that middle school students tend to use oppositional comments more than co-construction of knowledge. Previous research findings about use of oppositional comments and co-construction of knowledge are similar to current study. For example, Sampson and Clark (2011) reported that students producing higher quality arguments challenge each other by using oppositional comments, but students producing low quality arguments use co-construction of knowledge more.

As assertions derived from similarities of two classes formed, one more assertion was proposed based on the differences between two classes. In this study, class 2 engaged in argumentation more than class 1. Assertion derived from differences of classes shows that "While more active class tend to engage in argumentation more in the weeks when they conducted experiments, less active class tend to engage in argumentation process more in the weeks when students use evidence cards to construct arguments." Accordingly, more active class frequently used expositional comment, oppositional comment and information seeking in the weeks when they conducted experiment. On the other hand, less active class frequently used expositional comment, oppositional comment and information seeking in the weeks when they did not conduct experiment (Less active class was active in the weeks they used evidence cards). In this study, four of six weeks included conducting experiments and two of weeks did not include experiments. Therefore, this study's focus on conducting experiment might make class 2 (i.e. the class which was active in experiment weeks) more active class and make class 1 (i.e. the class which was not active in the weeks that experiment was conducted) less active class. According to this result, class 2 (i.e. more active class) focused on first hand data which they produced in experiments, but they less emphasized the second hand data given in evidence cards. Wallace (2004) categorized such students who focus on only first-hand data as first-hand observers. These students have positivist epistemology and they think that scientific knowledge is only obtained from hand on activity data. First hand observers could not integrate class explanation and textbook explanation. Their explanations are limited with first-hand data and not supported by theories (Wallace, 2004). On the other hand, class 1 (i.e. less active class) focused on second hand data which was obtained from evidence cards, but they did not use first hand data coming from experiments. Such kind of students focusing on second hand data, but ignoring first hand data can be categorized as reader (Wallace, 2004). These students focus on facts, concepts, principle and explanations, but they ignore hands on activities. They think that first hand data does not include theoretical explanation. These students oversimplify role of evidence and observation in science. These students also think that laboratory work is done to verify knowledge and correct knowledge exists in second hand data (Wallace, 2004). As it is seen, readers' (less active class) epistemology is close to positivist epistemology too. In conclusion, this assertion suggests that although classes differed in terms of their engagement to argumentation process, none of the class focused on using both first hand data and second hand data and their epistemology are close to positivist epistemology which is not sophisticated.

Overall, as three important aspect of science education which are scientific literacy, argumentation schemes, and engagement in argumentation process are thought together; findings might become more meaningful. First of all, this study aimed to increase students' scientific literacy. However, only content knowledge improved among three aspects of scientific literacy. On the other hand, epistemological beliefs and science process skills did not change. Students' engagement in argumentation process can be related with improvement of content knowledge. Accordingly, when students knew the content, they used oppositional comments. Using oppositional comments, students persuaded each other and so persuaded ones might have improved their content knowledge. On the other hand, when students did not know the content, they used co-construction of knowledge. Use of co-construction of knowledge sometimes did not let students to reach scientific knowledge. For example, students assisted each other to understand the concepts, events and theories, but they reached some wrong ideas and explanation through co-operation. When such events occurred, teacher added some scientific explanations and tried to persuade students regarding elimination of wrong ideas. Similarly, use of expositional comments might have caused exhibition of wrong ideas. In some cases, students proposed their ideas as expositional comments starting argumentation process. However, these initial ideas revealed as expositional comments including wrong knowledge were sometimes not examined by others. On the other hand, when these expositional comments were criticized by others, owners of expositional comments changed their initial ideas and they reached scientific knowledge through persuasion. By this way, their content knowledge might have improved.

Regarding argumentation schemes and content knowledge, it can not be claimed that specific use of one argumentation schemes leads scientific knowledge. At least, no clue for this claim was found in this study. For example, students frequently used argument from sign during argumentation. Some uses of argument from sign included scientific knowledge, but some others did not. Therefore, it can not be claimed that argument from sign leads to correct knowledge. Furthermore, argumentation schemes are presumptive which means that they have lack of evidence and they can change when new evidence emerges. When they have lack of evidence, it is not meaningful to claim that some specific argumentation schemes directly address scientific knowledge. Although specific use of argumentation schemes is not directly related with scientific knowledge, continuous use of different or the same argumentation schemes might cause reaching scientific knowledge. For example, one argument from sign might include some missing or wrong ideas because it does not include enough evidence to reach scientific knowledge. On the other hand, additional schemes regardless of their types can increase the probability of reaching scientific knowledge because each new scheme will bring some new evidence. Wrong ideas will replace with correct ones thanks to new evidence when new argumentation schemes emerge. At the end, it is expected that all wrong ideas asserted by argumentation schemes are eliminated and students reach scientific knowledge by adding all evidence each other through reasoning.

Although participants' content knowledge increased, their epistemological beliefs did not change. Participants' epistemological beliefs can also be related with argumentation schemes and students' engagement in argumentation process. More specifically, this study showed that students tended to use either argument from expert opinion when they got second hand data or argument from evidence to hypothesis when they used first hand data. However, they did not integrate these two different data sources and they did not use argument from evidence to hypothesis and argument from expert opinion together. Their ignorance of one type of data when they focused on another data source shows that students did not act like scientists. For example; scientists conduct their investigations to produce first hand data. After that scientists feed their findings with theories and laws found in second hand data. However, students did not feed their first hand data with second hand data suggesting that they did not have sophisticated epistemological beliefs although quantitative findings suggested students had sophisticated beliefs. Students' less sophisticated epistemological beliefs might cause them not to engage in argumentation so much and not engaging argumentation might cause no change on students' epistemological beliefs as feedback loop. Regarding the link between epistemological beliefs and students' engagement in argumentation process, students' engagement in argumentation process also supported the idea that participants held less sophisticated epistemological beliefs. For example, students mostly used expositional comments in this study. When students only used expositional comments and do not engage in argumentation, and not criticize the ideas, students might think that the knowledge they proposed is certain and does not change. These ideas about certainty and no change of scientific knowledge is indicator of having less sophisticated epistemological beliefs. Likewise, in this study, one of the observed classes actively engaged in argumentation when experiments were conducted and other class actively engaged in argumentation when evidence cards were provided them. The class which was active during experiment weeks focused on first hand data, but this class ignored the theories, laws and concepts found in second hand data. These students might have thought that scientific knowledge is only obtained from first hand data reflecting less sophisticated beliefs (Wallace, 2004). Simlarly, students who were active when evidence cards were given them might have thought that scientific knowledge found in second hand sources are certain and they mignt have ignored the role of observation and experiment in science. They might have thought that laboratory work is done to verify knowledge found in second hand sources. Such ideas like ‘Scientific knowledge is certain’ and ‘Experiment is done to verify knowledge’ are also indicator of having less sophisticated epistemological beliefs.

In this study, students’ science process skills also did not improve. One reason why participants SPS did not improve can be the number activities. In this study, students conducted one main activity for each week so students conducted six main activities in total, and four of these activities included experiments. It is possible that four experiment activity in total did not enough to detect a statistically significant improvement in students’ science process skills. On the other hand, when the qualitative data were examined it was realized that students actively exhibit their process skills through argumentation schemes and their engagement in argumentation process to some extent. The argumentation schemes that students used their science process skills were argument from sign, argument from evidence to hypothesis, and argument from comparison in this study. Accordingly, students used their observation and inference skills when they used argument from sign. This scheme was frequently used in the study, hence, it can be claimed that students were good at these two basic skills. Likewise, students used argument from evidence to hypothesis. Argument from evidence to hypothesis was mainly used when students falsify or verify their hypothesis via experiments. Use of this scheme might show that students reached some level in formulating hypothesis, designing experiment and identifying variables skills. Similarly, students frequently used argument from comparison. This schema was used when students explained their experiment and control group. In other word, students exhibit their controlling and identifying variable skills using argument from comparison. Although students actively used these three schemes showing their science process skills, there were some other schemes related with science process skills; however, these schemes were rarely used in current study. For example, argument from verbal classification is used when people define something with their boundaries. This scheme can be used when students operationally define their variables. However, this scheme was used only once in this study. Hence, this finding may suggest that students had difficulty in defining operationally skills. Likewise, argument from corelation to cause is another scheme which is related with identifying variables. When students relate two variables each other, they use argument from corelation to cause. However, students in this study did not use argument from corelation to cause frequently. Due to fact that students did not frequently use argument from corelation to cause, it is possible that students did not reinforce their identifying variable skills and so their identifying skills might not have improved. In conclusion, if students had used argument from verbal classification and argument from corelation to cause more in this study, this might have contributed to their science process skills.

Likewise, there were some traces between students' engagement in argumentation process and science process skills. Firstly, this study showed that students mainly used information seeking component when students conduct experiment. Therefore, this component can be related with science process skills. For example, groups asked the details of experimental designs to presenting group and presenting group explained their experimental designs showing their designing experiment skills. Similarly, oppositional comments were used when presenting group's hypothesis and results conflicted. When students discussed falsification of hypothesis, they mentioned formulating hypothesis skill. Besides, presenting group's design of the experiment was faulty in some cases. When such cases were detected, other groups used oppositional comments to show presenting group's experiment design was inconvenient. Hence, students' use of oppositional comments might have fed their designing experiment skills. Similarly, co-construction of knowledge might have supported students' communication skill because students cooperatively reached knowledge when they used co-construction of knowledge. Lastly, students used expositional comments by proposing their ideas. Some of these ideas revealed clues regarding students' science process skills. In this expositional comments; students explained their hypothesis, variables, and experiment designs. Hence, students showed their formulating hypothesis, identifying variables, controlling variables, designing experiment skills by the use of expositional comments. Their expositional comments were correct in general. However, small group discussions were not analyzed in this study. Therefore, we do not know which group member contributed to the preparation of expositional comments. Such expositional comments can be product of all group members' effort or only one or two members' effort. It is possible that some group members did not contribute on preparation of expositional comments and so their science process skills might not improve in this study, but this is only speculation in this point because small group discussions were not analyzed in this study as explained above.

Although argumentation schemes and students' engagement in argumentation process are related with scientific literacy components including content knowledge, epistemological beliefs, and science process skills; no clear pattern was observed between argumentation schemes (e.g., argument from sign) and students' engagement in argumentation process components (e.g. expositional comments) except the consistency between information seeking component of engagement and argument from position to know scheme. Both information seeking component and argument from position to know emerged as questions when students needed additional knowledge. Hence, the link between only these two themes was established.

To sum up, results of current study were discussed comparing the results of previous studies. Implications of the current study regarding ABI effect on scientific literacy components (e.g. content knowledge), argumentation schemes and students' engagement in argumentation process are presented in next part.

## 5.4. Implications

Current study provides some implications. Most of the implication addresses ABI research and so these implications are for researchers. Besides, the study has some implications for curriculum developers and science teachers.

First three implications of the study are about epistemological beliefs. In this study, epistemological beliefs were measured by quantitative scale. However, the study showed that quantitative findings about epistemological beliefs conflicted with qualitative data. While quantitative data suggested that students had constructivist epistemology, qualitative data showed that students had positivist epistemology. Therefore, ABI research focusing on epistemological beliefs should triangulate qualitative and quantitative data in future research. Next, participants' epistemological beliefs did not improve in this study. Explicit epistemology instruction might be helpful to improve participants' epistemological beliefs in ABI research. For example, students might discuss epistemology components specifically when they engage in activities in argumentation. When they discuss epistemology components in argumentation, they might be aware of their epistemological status and they can link epistemology and argumentation. Similarly, students actively used evidence cards in this study, but they did not criticize evidence cards. This is evidence that students did not have sophisticated epistemological beliefs. Therefore, students should be informed that evidence cards can also be critically evaluated.

This study also showed that there was no significant change in students' science process skills in ABI treatment. Therefore, explicit science process skill instruction can be added to following ABI studies to improve participants' SPS. In this SPS instruction, students can be alerted about their science process skills when they conduct investigation and they can improve SPS through practice in ABI research.

In this study, there are also implications about argumentation schemes students used in ABI treatment. Findings of the study showed that total frequency of argumentation schemes did not change over time. On the other hand, total type of argumentation schemes increased throughout the study. Researchers studying on reasoning can focus on participants’ use of total type of argumentation schemes instead of total number of frequency of argumentation schemes. Increase of total type of argumentation schemes can be indicator of development of participants’ reasoning in ABI treatment because appealing to different schemes means thinking from different perspectives which require in depth reasoning. On the other hand, focusing on frequency of total number of argumentation schemes might mislead researchers in terms of participants’ reasoning. For example, participants may increase frequency of total number of argumentation schemes, but they can use limited type of argumentation schemes showing limited reasoning.

This study also showed that middle school students mainly use five argumentation schemes which are argument from expert opinion, argument from sign, argument from position to know, argument from comparison, and argument from evidence to hypothesis although there are 25 argumentation schemes in total. While researchers prepare argumentation activities, they can consider these five argumentation schemes that students mostly use. By this way, students can better and faster adapt to argumentation process. On the other hand, if activities are not linked with these most used schemes, students may have difficulty in getting familiarity to argumentation and this cause the decline of efficiency of ABI treatment. Furthermore, researchers and teachers can design other activities supporting students’ use of other argumentation schemes which were used less frequently in this study (e.g. argument from corelation to cause). For example, teachers may ask students the relationship between different variables. When students focus on the potential relationships, they probably will use argument from corelation to cause and argument from cause to effect more frequently. Moreover, teachers and researchers may provide more than one solution to a problem and students are expected to select one of the solutions. When students comparing possible solutions, they can decide which solution is good or bad. By this way, students can frequently use argument from consequences which was not used so much in this study.

The study also showed that participants’ use of argumentation schemes changes depending on the data sources. Accordingly, students tend to use argument from evidence to hypothesis when first hand data are provided, but they use argument from expert opinion when second hand data are provided. However, they ignored use of argument from expert opinion when second hand data were not given them explicitly. Likewise, students ignored use of argument from evidence to hypothesis when they did not conduct investigation in which first hand data is obtained from. Therefore, researchers can provide students both first hand data and second hand data and students should be explicitly informed that they are expected to use both data sources and integrate them. By this way, students can integrate their observations obtained from first hand data (e.g. experiments) and theoretical knowledge obtained from second hand data (e.g. evidence cards, textbooks, and teacher), so they can learn better.

Similarly, the study has implications for students' engagement in argumentation process. Findings of the study showed that students do not tend to engage in argumentation process when the activities are unexciting and similar. Therefore, researchers should select surprising activities which motivate students to engage in argumentation. Using exciting activities seems to be first condition for effective engagement in argumentation. Likewise, researchers should not create idealized conditions which eliminate the error and alternative results and alternative ideas. For example, science experiment was replaced with simulation activity in the last week of this study. By this way, problematic real life conditions replaced with idealized conditions provided by simulation. Removing problematic conditions and adding idealized conditions caused converging ideas of students and students did not engage in argumentation frequently in this week. Therefore, it is suggested that faulty experiments can be good for argumentation and researchers should follow the faulty experiments, their problems, different ideas explaining the problem and alternative ideas to solve the problem instead of removing faulty experiments and using simulations providing idealized conditions which do not require any discussion. Although faulty experiments results in more engagement in argumentation process, students may not reach correct content knowledge. For example, students could not observe inverse relationship between wire resistance and length of wire in pilot study and therefore, they could not reach the content knowledge including the inverse relationship between wire length and wire resistance. After students engage in whole class discussion and do not reach correct knowledge, teachers can provide second hand data like textbook as evidence cards explaining reverse relationship between wire resistance and wire length. By this way, students' content knowledge can increase. Likewise, simulation programs can be shown to students as additional activity after students conduct their faulty experiments. Students can be asked to compare their faulty experiment results and simulation results and this comparison can increase their content knowledge. Similarly, teachers can use analogies and visual representation that concrete abstract relations (e.g. wire resistance and wire length relation) after whole class discussion ends. Teachers' guidance and additional explanations about content after discussion can assist students to understand curricular objectives including content knowledge.

Likewise, researchers should determine their major outcome prior to the study and they should act based on this outcome, and researchers should not change their outcome throughout the study. The major outcome for the researcher can be defined as the outcomes researchers expect from participants as a result of participants' engagement in argumentation process. These outcomes can be reaching correct content knowledge, increasing skills (e.g. argumentation skills), producing high quality argument, examining students' engagement in argumentation process. For example, major outcome of this study is examining students' engagement in argumentation process in this study (i.e., main study). However, reaching correct knowledge was another outcome of the study (i.e., pilot study). While adjustments were done in pilot study, reaching correct knowledge was considered instead of major outcome (i.e. engaging in argumentation process). This situation caused problems in main study regarding students' engagement in argumentation. Accordingly, experiment was removed from the main study and simulation activity was added for the last week of main study considering correct knowledge as outcome when students could not reach the correct knowledge through experiment in pilot study. However, when simulation activity was added to main study instead of experiment, students did not engage in argumentation which is the main outcome in last week of main study. Therefore, researchers are advised to make adjustments in their studies considering their major outcome. Otherwise, their adjustments done in pilot study may not be efficient for the outcomes of main study.

Next, this study showed that students’ engagement in argumentation process did not improve throughout the study. One of the possible reasons can be nature of activities. For example, unexciting activities, idealized conditions, and use of different data sources probably affected students’ engagement in argumentation, so no increasing or decreasing trend was observed throughout the study regarding students’ engagement in argumentation process. Therefore, researchers are advised not to use unexciting activities and activities including idealized conditions in argumentation studies. If such activities which decrease students’ engagement in argumentation process are not used in argumentation studies, it is possible that students engage in argumentation more because surprising activities will increase their interest to engage in argumentation. When students get familiarity to the argumentation process over time, it is possible that there will be an increasing trend regarding students’ engagement in argumentation process.

Current study also has some implications regarding the components of students’ engagement in argumentation. Previous research showed that oppositional comment is a characteristic of argumentation and it is observed more than others when students engaged in argumentation (Sampson & Clark, 2011). However, expositional comment was used more than others in this study. On the other hand, previous research (e.g. Sampson & Clark, 2011) did not focus on using expositional comment in argumentation. Therefore, limited knowledge has been known about expositional comment. Previous research emphasized on oppositional comments and co-construction of knowledge because these two components assist students to reach content knowledge. On the other hand, current study showed that using expositional comments are also important during argumentation because students propose their ideas using expositional comments and these expositional comments starts argumentation. Other engagement components like oppositional comments and co-construction of knowledge are constructed based on the ideas proposed as expositional comments. In other word, expositional comments provide data and ideas for other engagement components. The more expositional comments are used as source for discussion, the deeper argumentation can students engage in. Therefore, researchers studying on students' engagement in argumentation can focus expositional comments as data source for argumentation and increase our understanding about this oversimplified, but the most common component.

Likewise, this study showed that when evidence cards are given to students in topic that students did not have prior knowledge, students mostly use expositional comments by proposing their ideas using evidence cards as expert opinion, but students did not use other three engagement components. This situation might decrease the quality of argumentation because questions are not asked, different ideas are not discussed and knowledge is not constructed in cooperation. Therefore, it can be better if researchers do not provide evidence cards to students when they do not know the content. Other activities which are discrepant, surprising, mystical and design based feeding their knowledge can be better for their engagement in argumentation.

Findings about oppositional comments can also be used in ABI research. This study showed that students use oppositional comments most when students have prior knowledge about topic and evidence cards are given them. Previous research showed that getting familiar to the argumentation research is difficult task and short period of argumentation studies cause students not to benefit from argumentation studies. This finding about oppositional comment might be helpful for students to get familiar to argumentation in a short time. Accordingly, some classes cannot engage in argumentation in initial weeks because of lack of familiarity as in the case of Class 1 in this study. However, if researchers choose a topic that students are knowledgeable and provide students evidence cards supporting their knowledge, students might frequently use oppositional comments even at the beginning of argumentation study and they may easier and earlier adapt to argumentation. This suggestion especially can be better for argumentation studies including limited time.

As little is known about expositional comments in argumentation process, there is little known about information seeking component of students' engagement in argumentation. Information seeking was not seen as important component for engagement in argumentation; however, current study showed that information seeking is important especially for revealing divergent points in argumentation. Therefore, frequent use of information seeking can be sign of deep discussion, and studying on information seeking can be helpful to develop our understandings about students' engagement in argumentation process. Similarly, science teachers are suggested to motivate students to ask questions to their peers throughout the argumentation process because asking questions to others (i.e. information seeking) can trigger and deepen discussion.

This study showed that students can easily use expositional comments because proposing their own ideas is easier than considering others' views. The study also showed that middle school students can frequently use oppositional comments. However, students cannot use co-construction of knowledge frequently. Even though they rarely use co-construction of knowledge, they may not reach to scientifically correct knowledge. This study showed that participants use co-construction of knowledge in three different conditions. Co-construction of knowledge was used when groups need assistance, groups reach same conclusion, and there is a lack of knowledge to explain observed phenomena. Therefore, researchers can set research design considering these three conditions. If researchers set such designs, participants most probably use co-construction of knowledge much more. Moreover, the study showed that students' use of co-construction of knowledge sometimes resulted in wrong content knowledge. In such situations, teachers and researchers can provide students evidence cards and learning activities assisting students to eliminate wrong content knowledge.

The study has also three implications for curriculum makers. In this study, students designed, prepared and conducted their own investigations in four of six weeks. However, students were asked to modify their research questions when research questions did not match with curricular objectives because teachers have to teach specific curricular objectives in Turkey. These curricular obligations might decrease students' motivations to participate in experiments because students might feel that they could not conduct their own experiments when they change the research question. In this point, curriculum makers are suggested not to focus on teaching specific objectives and teachers should have autonomy in teaching. For example, curriculum makers can determine important core ideas like heat and electricity without emphasizing specific objectives. Then, teachers can be free when teaching core ideas. If these two suggestions are done, students can follow their own research questions freely when they conduct investigations and they can learn the core ideas. Another implication for curriculum makers is about materials. In this study, laboratory materials and simple materials were used to conduct experiments. Curriculum makers can provide more and contemporary grade specific (i.e. middle school) materials because students' creativity and imagination have no limit. In this study, students produced interesting ideas and research questions throughout the study. If students' imaginations are supported by additional material, it is possible that students can think, make reasoning, argue and learn more. Next, this study was conducted in science units which included objectives about experiments and discussions. Therefore, it can be said that ABI research aligns with objectives having experiments and discussions. In this point, curriculum developers can add experiment and discussion objectives to other science topics found in middle school curriculum. If this is done, researchers can conduct ABI research on other science topics because researchers have to follow curricular objectives in Turkey in their research. Conducting ABI research on different science topics in middle school level can further increase our understanding about ABI research.

The study provides further recommendations derived from limitations of this study. Firstly, qualitative part of the study focused on whole class discussion for students' engagement in argumentation process, but pre-activity discussions and small group discussions were not used for analysing students' engagement in argumentation process. Future studies can focus both pre-activity, small group and whole class discussions to better understand nature of middle school students' engagement in argumentation. Secondly, short period of ABI treatment is another issue for this study. Previous studies claimed that if duration of treatment increases, participants benefit from ABI more. This study lasted six weeks and future studies can last longer than this period to increase treatment efficiency.

## REFERENCES


Akben, N. (2015). Improving science process skills in science and technology course activities using the inquiry method. *Egitim ve Bilim*, *40*(179), 111-132.

Akkus, R., Gunel, 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.

Appleton, K. (2007). Elementary science teaching (Chapter 18). In S.K.Abell and N.G. Lederman (Eds.), *Research on Science Teacher Education*, pp. 493-535, New York: Routledge.

Archila, P. A. (2015). Using History and Philosophy of Science to Promote Students' Argumentation. *Science & Education*, *24*(9-10), 1201-1226.

Aslan, S. (2016). Argümantasyona dayalı laboratuar uygulamaları: Bilimsel süreç becerilerine ve laboratuar dersine yönelik tutuma etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)*, *31*(4), 762-777.

Aydeniz, M., Pabuccu, A., Cetin, P. S., & Kaya, E. (2012). Argumentation and students'conceptual understanding of properties and behaviors of gases. *International Journal of Science and Mathematics Education*, *10*(6), 1303-1324.

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. *Cognitive domain.*

Bunterm, T., Lee, K., Ng Lan Kong, J., Srikoon, S., Vangpoomyai, P., Rattanavongsa, J., & Rachahoon, G. (2014). Do different levels of inquiry lead to different learning outcomes? A comparison between guided and structured inquiry. *International Journal of Science Education, 36*(12), 1937-1959.

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, *22*(2), 169-177.

Can, B. (2008). *İlköğretim öğrencilerinin bilimin doğasi ile ilgili anlayışlarını etkileyen faktörler*. (Doctoral dissertation). Dokuz Eylül Üniversitesi Eğitim Bilimleri Enstitüsü.

Cansiz, M. (2014). *The effect of history of science instruction on elementary students' scientific literacy.* (Unpublished doctoral dissertation). Middle East Technical University, Ankara, Turkey.

Carlson, J., & Daehler, K. R. (2019). The refined consensus model of pedagogical content knowledge in science education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 77–92). Springer Singapore.

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K–12 science contexts. *Review of Educational Research, 80*(3), 336-371.

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

Cetin, P. S., Dogan, N., & Kutluca, A. Y. (2014). The quality of pre-service science teachers' argumentation: Influence of content knowledge. *Journal of Science Teacher Education*, *25*(3), 309-331.

Cetin, P. S., Metin, D., Capkinoglu, E., & Leblebicioglu, G. (2016). Seeking the trace of argumentation in Turkish science curriculum. *Science Education International, 27*(4), 570-591.

Chang, YH, Chang, CY, & Tseng, YH. (2010). Trends of science education research: an automatic content analysis. *Journal of Science Education and Technology, 19*, 315–331.

Chen, Y.C. (2011). *Examining the integration of talk and writing for student knowledge construction through argumentation*. (Doctoral dissertation). University of Iowa.

Chen, Y. C., Hand, B., & McDowell, L. (2013). The effects of writing-to-learn activities on elementary students' conceptual understanding: Learning about force and motion through writing to older peers. *Science Education*, *97*(5), 745-771.

Chen, Y. C., Hand, B., & Park, S. (2016). Examining elementary students' development of oral and written argumentation practices through argument-based inquiry. *Science & Education*, *25*(3-4), 277-320.

Chen, H. T., Wang, H. H., Lu, Y. Y., Lin, H. S., & Hong, Z. R. (2016). Using a modified argument-driven inquiry to promote elementary school students' engagement in learning science and argumentation. *International Journal of Science Education*, *38*(2), 170-191.

Chin, C., & Osborne, J. (2010). Supporting argumentation through students' questions: Case studies in science classrooms. *Journal of the Learning Sciences*, *19*(2), 230-284.

Choi, A., Hand, B., & Norton-Meier, L. (2014). Grade 5 students' online argumentation about their in-class inquiry investigations. *Research in Science Education, 44*(2), 267-287.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Lawrence Earlbaum Associates.

Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, *29*(2), 186-204.

Creswell, J.W. (2006). *Qualitative inquiry and research design: Choosing among five approaches (2nd ed.).* Thousand Oaks,. CA: Sage.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches (3rd ed.).* Sage Publications, Inc

Cronje, R., Murray, K., Rohlinger, S., & Wellnitz, T. (2013). Using the science writing heuristic to improve undergraduate writing in biology. *International Journal of Science Education, 35*(16), 2718-2731.

Dawson, V., & Venville, G. (2013). Introducing high school biology students to argumentation about socioscientific issues. *Canadian Journal of Science, Mathematics and Technology Education*, *13*(4), 356-372.

Demirbağ, M., & Günel, M. (2014). Argümantasyon tabanlı fen eğitimi sürecine modsal betimleme entegrasyonunun akademik başarı, argüman kurma ve yazma becerilerine etkisi. *Kuram Ve Uygulamada Eğitim Bilimleri, 14*(1), 373-392.

Dindar, H., & Taneri, A. (2011). MEB'in 1968, 1992, 2000 ve 2004 yıllarında geliştirdiği fen programlarının amaç, kavram ve etkinlik yönünden karşılaştırılması. *Kastamonu Eğitim Dergisi*, *19*(2), 363-378.

Dunkhase, J. A. (2003). The coupled-inquiry cycle: A teacher concerns-based model for effective student inquiry. *Science Educator*, *12*(1), 10-15.

Durmaz, H., & Mutlu, S. (2017). The effect of an instructional intervention on elementary students' science process skills. *The Journal of Educational Research, 110*(4), 433-445.

Duschl, R. A. (2007). Quality argumentation and epistemic criteria. In S. Erduran & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom-based research* (pp. 159–175). Dordrecht: Springer.

Ebel, R. L., & Frisbie, D. A. (1991). *Essentials of educational measurement (5th ed).* Englewoods Cliffs.

Elder, A.D. (2002). Characterizing fifth grade students' epistemological beliefs in science. In P.R. Pintrich (Ed.), *Personal epistemology: The psychology of beliefs about knowledge and knowing*, (pp. 347-364). Lawrence Erlbaum Associates, Mahwah, NJ, USA.

Emig, B. R., McDonald, S., Zembal-Saul, C., & Strauss, S. G. (2014). Inviting argument by analogy: analogical-mapping-based comparison activities as a scaffold for small-group argumentation. *Science Education, 98*(2), 243-268.

Erduran, S. (2007). Methodological foundations in the study of argumentation in science classrooms. In S. Erduran, & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom based research* (pp. 47-69). Dordrecht, the Netherlands: Springer.

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.

Erkol, M., Kışoğlu, M., & Gül, Ş. (2017). The effect of argument-based science inquiry report format on pre-service science teachers' achievements and attitudes towards science laboratory [Argümantasyon tabanlı bilim öğrenme yaklaşımı rapor formatının öğretmen adaylarının başarılarına ve fen bilgisi laboratuvarına yönelik tutumlarına etkisi]. *İlköğretim Online*, *16*(2), 614-627.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education (8th ed.)*. New York: McGraw-Hill.

Garcia-Mila, M., & Andersen, C. (2007). Cognitive foundations of learning argumentation. In S. Erduran, & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom based research* (pp. 29-45). Dordrecht, the Netherlands: Springer.

Geban, Ö., Askar, P., & Özkan, İ. (1992). Effects of computer simulations and problem-solving approaches on high school students. *The Journal of Educational Research, 86*(1), 5-10.

Germann, P. J. (1989). Directed-inquiry approach to learning science process skills: treatment effects and aptitude-treatment interactions. *Journal of Research in Science Teaching, 26*(3), 237-250.

Gök, G. (2014). *The effect of 7E learning cycle instruction on 6th grade students' conceptual understanding of human body systems, self-regulation, scientific epistemological beliefs, and science process skills* (Unpublished doctoral dissertation). Middle East Technical University, Ankara, Turkey.

Gravetter, F. J., & Wallnau, L. B. (2007). *Statistics for the behavioral sciences*. Belmont, CA: Thomson Learning.

Greenbowe, T. J., Poock, J. R., Burke, K. A., & Hand, B. M. (2007). Using the science writing heuristic in the general chemistry laboratory to improve students' academic performance. *Journal of Chemical Education, 84*(8), 1371-1379.

Greenbowe, T. J., Rudd, J. A., & Hand, B. M. (2007). Using the science writing heuristic to improve students' understanding of general equilibrium. *Journal of Chemical Education, 84*(12), 2007-2011.

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.

Hahs-Vaughn, D. L. (2017). *Applied multivariate statistical concepts*. Taylor & Francis. New York: Routledge.

Hand, B., & Keys, C. W. (1999). Inquiry investigation: A new approach to laboratory reports. *The Science Teacher, 66*(4), 27–29.

Hand, B., Norton-Meier, L. A., Gunel, M., & Akkus, R. (2016). Aligning teaching to learning: A 3-year study examining the embedding of language and argumentation into elementary science classrooms. *International Journal of Science and Mathematics Education, 14*(5), 847-863.

Hand, B., Wallace, C. W., & Yang E. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: quantitative and qualitative aspects. *International Journal of Science Education, 26*(2), 131-149.

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

Heng, L. L., Surif, J., & Seng, C. H. (2015). Malaysian Students' Scientific Argumentation: Do groups perform better than individuals? *International Journal of Science Education, 37*(3), 505-528.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research, 67*(1), 88-140.

Hohenshell, L. M., & Hand, B. (2006). Writing-to-learn Strategies in Secondary School Cell Biology: A mixed method study. *International Journal of Science Education, 28*(2-3), 261-289.

Hong, Z. R., Lin, H. S., Wang, H. H., Chen, H. T., & Yang, K. K. (2013). Promoting and scaffolding elementary school students' attitudes toward science and argumentation through a science and society intervention. *International Journal of Science Education, 35*(10), 1625-1648.

Jimenez-Aleixandre, M. P. (2008). Designing argumentation learning environments. In S. Erduran, & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom based research* (pp. 91-115). Dordrecht, the Netherlands: Springer.

Jimenez-Aleixandre, M.P., & Erduran, S. (2008). Argumentation in science education: An overview. In S. Erduran, & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom based research* (pp. 3-25). Dordrecht, the Netherlands: Springer.

Jin, H., Hokayem, H., Wang, S., & Wei, X. (2016). A US-China interview study: Biology students' argumentation and explanation about energy consumption issues. *International Journal of Science and Mathematics Education*, 14(6), 1037-1057.

Jönsson, A. (2016). Student performance on argumentation task in the swedish national assessment in science. *International Journal of Science Education, 38*(11), 1825-1840.

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.

Khishfe, R. (2012). Relationship between nature of science understandings and argumentation skills: A role for counterargument and contextual factors. *Journal of Research in Science Teaching, 49*(4), 489-514.

Khishfe, R. (2014). Explicit nature of science and argumentation instruction in the context of socioscientific issues: An effect on student learning and transfer. *International Journal of Science Education*, *36*(6), 974-1016.

Khishfe, R., Alshaya, F. S., BouJaoude, S., Mansour, N., & Alrudiyan, K. I. (2017). Students' understandings of nature of science and their arguments in the context of four socio-scientific issues. *International Journal of Science Education, 39*(3), 299-334.

Kim, S., & Hand, B. (2015). An analysis of argumentation discourse patterns in elementary teachers' science classroom discussions. *Journal of Science Teacher Education, 26*(3), 221-236.

Kim, H., & Song, J. (2006). The features of peer argumentation in middle school students' scientific inquiry. *Research in Science Education, 36*(3), 211-233.

Kind, P. M., Kind, V., Hofstein, A., & Wilson, J. (2011). Peer Argumentation in the School Science Laboratory—Exploring effects of task features. *International Journal of Science Education, 33*(18), 2527-2558.

Kingir, S., Geban, O., & Gunel, M. (2012). How does the science writing heuristic approach affect students' performances of different academic achievement levels? A case for high school chemistry. *Chemistry Education Research and Practice, 13*(4), 428-436.

Kingir, S., Geban, O., & Gunel, M. (2013). Using the science writing heuristic approach to enhance student understanding in chemical change and mixture. *Research in Science Education, 43*(4), 1645-1663.

Koksal, E. A., & Berberoglu, G. (2014). The effect of guided-inquiry instruction on 6th grade Turkish students' achievement, science process skills, and attitudes toward science. *International Journal of Science Education, 36*(1), 66-78.

Kolsto, S. D., & Ratcliffe, M. (2007). Social aspects of argumentation. In S. Erduran, & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom based research* (pp. 117-136). Dordrecht, the Netherlands: Springer.

Konstantinidou, A., & Macagno, F. (2013). Understanding students’ reasoning: argumentation schemes as an interpretation method in science education. *Science & Education, 22*(5), 1069-1087.

Lawson, A. (2003). The nature and development of hypothetico-predictive argumentation with implications for science teaching. *International Journal of Science Education, 25*(11), 1387-1408.

Lederman, N. G., & Lederman, J. S. (2012). Nature of scientific knowledge and scientific inquiry: Building instructional capacity through professional development. In B.J. Fraser E, K.G. Tobin & C.J. McRobbie (Eds.), *Second international handbook of science education* (pp. 335-359). Springer, Dordrecht.

Lee, M. H., Wu, Y. T., & Tsai, C. C. (2009). Research trends in science education from 2003 to 2007: A content analysis of publications in selected journals. *International Journal of Science Education, 31*(15), 1999-2020.

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

Lin, T. J., Lin, T. C., Potvin, P., & Tsai, C. C. (2018). Research trends in science education from 2013 to 2017: A systematic content analysis of publications in selected journals. *International Journal of Science Education, 41*(3), 367-387.

Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Lincoln, Y. S., & Guba, E. G. (2000). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The handbook of qualitative research 2nd ed*. (pp. 163–188). Beverly Hills, CA: Sage.

Liu, S., & Roehrig, G. (2019). Exploring science teachers' argumentation and personal epistemology about global climate change. *Research in Science Education, 49*(1), 173-189.

Macagno, F., Mayweg-Paus, E., & Kuhn, D. (2015). Argumentation theory in education studies: Coding and improving students' argumentative strategies. *Topoi, 34*(2), 523-537.

Martin, A. M., & Hand, B. (2009). Factors affecting the implementation of argument in the elementary science classroom. A longitudinal case study. *Research in Science Education, 39*(1), 17-38.

McDonald, C. V. (2010). The influence of explicit nature of science and argumentation instruction on preservice primary teachers' views of nature of science. *Journal of Research in Science Teaching, 47*(9), 1137-1164.

McDonald, C. V., & McRobbie, C. J. (2012). Utilising argumentation to teach nature of science. In B.J. Fraser E, K.G. Tobin & C.J. McRobbie (Eds.), *Second international handbook of science education* (pp. 969-986). Springer, Dordrecht.

Mendonça, P. C. C., & Justi, R. (2014). An instrument for analyzing arguments produced in modeling-based chemistry lessons. *Journal of Research in Science Teaching, 51*(2), 192-218.

Merriam, S.B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* Sage.

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching.* Pearson.

Ministry of National Education (MoNE). (2013). *Elementary 6th, 7th and, 8th grades science curriculum.* Ankara, Turkey.

Ministry of National Education (MoNE). (2018). *Elementary 6th, 7th and, 8th grades science curriculum.* Ankara, Turkey.

Namdar, B., & Shen, J. (2016). Intersection of argumentation and the use of multiple representations in the context of socioscientific issues. *International Journal of Science Education, 38*(7), 1100-1132.

Naylor, S., Keogh, B., & Downing, B. (2007). Argumentation and primary science. *Research in Science Education*, *37*(1), 17-39.

Next Generation Science Standards (NGSS). (2013). *Crosscutting Concepts.* Retrieved from: http://www.nextgenscience.org/sites/ngss/files/Appendix%20G%20-%20Crosscutting%20Concepts%20FINAL%20edited%204.10.13.pdf

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.

Özkan, Ş. (2008). *Modeling elementary students' science achievement: the interrelationships among epistemological beliefs, learning approaches, and self-regulated learning strategies* (Unpublished Doctoral Dissertation). Middle East Technical University, Ankara.

Padilla, M. J. (1990). The science process skills. *Research Matters-to the science teacher,* 9004.

Pallant, J. F. (2007). *SPSS survival manual: A step-by-step guide to data analysis with SPSS*. New York, NY: McGraw-Hill.

Pella, M. O., O'hearn, G. T., & Gale, C. W. (1966). Referents to scientific literacy. *Journal of Research in Science Teaching, 4*(3), 199-208.

Perry, B., & Dockett, S. (1998). Play, argumentation and social constructivism. *Early Child Development and Care, 140*(1), 5-15.

Poock, J. R., Burke, K. A., Greenbowe, T. J., & Hand, B. M. (2007). Using the science writing heuristic in the general chemistry laboratory to improve students' academic performance. *Journal of Chemical Education, 84*(8), 1371-1379.

Popham, W. J. (2005). *Assessment for educational leaders*. Pearson/Allyn and Bacon.

Roberts, D. A. (2007). Scientific literacy/ Science literacy. In S. K. Abell, & N. G. Lederman, *Handbook of research on science education*, (pp. 729-780). Routledge, New York.

Rudd, J. A., Greenbowe, T. J., Hand, B. M., & Legg, M. J. (2001). Using the science writing heuristic to move toward an inquiry-based laboratory curriculum: An example from physical equilibrium. *Journal of Chemical education, 78*(12), 1680.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 41*(5), 513-536.

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

Sampson, V., & Clark, D. B. (2006). Assessment of argument in science education: a critical review of the literature. In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.). *Proceedings of the Seventh International Conference of the Learning Sciences—Making a Difference* (pp. 655–661). Mahwah, NJ: Erlbaum.

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

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

Sampson, V., Enderle, P. J., & Walker, J. P. (2012). The development and validation of the assessment of scientific argumentation in the classroom (ASAC) observation protocole: A tool for evaluating how students participate in scientific argumentation. In M. S. Khine (Ed). *Perspectives on scientific argumentation: Theory, practice and research* (pp. 235-272). Springer.

Sanderson, B. A., & Kratochvil, D. W. (1971). *Science, a Process Approach*. American Institute for Research in the Behavioral Science.

Sandoval, W. A. (2003). Conceptual and epistemic aspects of students' scientific explanations. *The Journal of the Learning Sciences, 12*(1), 5-51.

Sandoval, W. A., & Millwood, K. A. (2008). What can argumentation tell us about epistemology? In S. Erduran, & M.P. Jimenez-Aleixandre (Eds.), *Argumentation in science education: Perspectives from classroom based research* (pp. 71-88). Dordrecht, the Netherlands: Springer.

Sawada, D., Piburn, M., Falconer, K., Turley, J., Benford, R., & Bloom, I. (2000). *Reformed teaching observation protocol (RTOP)* (Tech. Rep No. IN00-1). Tempe, AZ: Arizona State University, Arizona Collaborative for Excellence in the Preparation of Teachers.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*(3), 498-504.

Schraw, G., Bendixen, L. D., & Dunkle, M. E. (2002). *Development and validation of the Epistemic Belief Inventory (EBI).* In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (p. 261–275). Lawrence Erlbaum Associates Publishers.

Shamuganathan, S., & Karpudewan, M. (2017). Science writing heuristics embedded in green chemistry: a tool to nurture environmental literacy among pre-university students. *Chemistry Education Research and Practice, 18*(2), 386-396.

Tabachnick, B. G., & Fidell, L. S. (2012). *Using multivariate statistics (6th Ed).* Boston, MA: Pearson.

Taylor, J. C., Tseng, C. M., Murillo, A., Therrien, W., & Hand, B. (2018). Using Argument-Based Science Inquiry to Improve Science Achievement for Students with Disabilities in Inclusive Classrooms. *Journal of Science Education for Students with Disabilities, 21*(1), 1-14.

Toulmin, S. E. (1958). *The philosophy of science* (Vol. 14). Genesis Publishing Pvt Ltd.

Tsai, C.-C., & Wen, L.M.C. (2005). Research and trends in science education from 1998 to 2002: A content analysis of publication in selected journals. *International Journal of Science Education, 27*, 3–14.

Tucel, S. T. (2016). *Exploring the effects of science writing heuristic (swh) approach on the eighth grade students' achievement, metacognition and epistemological beliefs.* (Unpublished master dissertation). Middle East Technical University, Ankara, Turkey.

Tucel, T. Cakiroglu, J. & Öztekin, C. (2015). Exploring the Effects of Science Writing Heuristic Approach on 8th Grade Students' Metacognition. European Science Education Research Association Conference (ESERA) August 31-Sept. 4, Helsinki, Finland (PP. 483-488)

Vieira, R. D., Bernardo, J. R. D. R., Evagorou, M., & Melo, V. F. D. (2015). Argumentation in Science Teacher Education: The simulated jury as a resource for teaching and learning. *International Journal of Science Education, 37*(7), 1113-1139.

Walker, J. P., & Sampson, V. (2013). Learning to Argue and Arguing to Learn: Argument-Driven Inquiry as a Way to Help Undergraduate Chemistry Students Learn How to Construct Arguments and Engage in Argumentation During a Laboratory Course. *Journal of Research in Science Teaching, 50*(5), 561-596.

Walker, J. P., Sampson, V., & Zimmerman, C. O. (2011). Argument-driven inquiry: An introduction to a new instructional model for use in undergraduate chemistry labs. *Journal of Chemical Education, 88*(8), 1048-1056.

Wallace, C. S. (2004). An illumination of the roles of hands-on activities, discussion, text reading, and writing in constructing biology knowledge in seventh grade. *School Science and Mathematics*, *104*(2), 70-78.

Walsh, E. M., & McGowan, V. C. (2017). 'Let your data tell a story:'climate change experts and students navigating disciplinary argumentation in the classroom. *International Journal of Science Education*, *39*(1), 20-43.

Walton, D. N. (1996). *Argumentation schemes for presumptive reasoning*. Psychology Press.

Walton, D., Reed, C., & Macagno, F. (2008). *Argumentation schemes*. Cambridge University Press.

Wojdak, J. M. (2010). An attention-grabbing approach to introducing students to argumentation in science. *Bioscience Education*, *15*(1), 1-3.

Yaman, F. (2019). Öğrencilerin sanal kimya laboratuvarı kullanarak hazırladıkları argümantasyona dayalı yazma etkinliklerinin çoklu gösterimler açısından incelenmesi. *Elementary Education Online, 18*(1), 207-225.

Yeşildağ-Hasançebi, F., & Günel, M. (2013). Effects of Argumentation Based Inquiry approach on disadvantaged students' science achievement. *Elementary Education Online, 12*(4), 1056-1073.

Yıldırım, A. (2012). *Effect of guided inquiry experiments on the acquisition of science process skills, achievement and differentiation of conceptual structure*. (Unpublished master dissertation). Middle East Technical University, Ankara, Turkey.

Zemplén, G. Á. (2011). History of science and argumentation in science education: Joining forces?. In *Adapting historical knowledge production to the classroom* (pp. 129-140). Brill Sense.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *39*(1), 35-62.

# APPENDICES

## A . SWH LABORATORY REPORT

Deneyin Adı: _______________________________ Adı Soyadı: __________

Deney Masası: _______________________________ Tarih: _______________

1. Başlangıç düşünceleri... Soru ya da sorularım nelerdir?

(Yani, bu konu/deney ile ilgili neleri merak ediyorum?)

2. Test... Sorularıma cevap bulmak için ne yaptım?

(Yani, merak ettiklerime ulaşmak için ne yaptım?)

3. Gözlemler ve bulgular... Yaptıklarım sonucunda neler buldum?

(Yani, merak ettiklerime ulaşmaya çalışırken bulduklarım ve gözlediklerim nelerdir?)

4. İddialar... Bulduklarım ve gözlediklerim sonunda ne iddia ediyorum?

(Yani, merak edip araştırdıklarım ile ilgili bu deney sonunda vardığım genel kanaatim kısa ve öz olarak...)

5. Deliller (Kanıtlar)...Bulduklarım ve gözlediklerim sonunda yukarıdaki iddiamı yaptım çünkü delillerim şunlardır: (Yani, bulduklarım ve gözlemlerimden ortaya çıkardığım iddiamı destekleyen deliller...)

6. Okuma ve karşılaştırmalar... Düşüncelerimin başkaları ile karşılaştırılması...

(Yani, düşüncemi arkadaşlarımın düşünceleri ile ve kitaptan okuduklarımla karşılaştırdım ve vardığım sonuç...)

7. Yansımalar... Düşüncelerim süreç içinde nasıl değişti?

(Yani, konu ile ilgili deneyin başındaki düşüncelerimle deneyin sonundaki düşüncelerimi karşılaştırarak değişimim ile ilgili vardığım sonuç...)

## B. TEACHER TEMPLATE

### 1. Öğrencilerin ön bilgilerinin ortaya çıkarılması

Günlük hayat ile ilgili bağlantılar yapılarak öğrencilerin ön bilgileri açığa çıkartılır. Bu aşamada öğretmen sorular sorarak sürekli öğrenciler ile iletişim halindedir. Öğrenci ön bilgileri ortaya konduktan sonra laboratuvar öncesi etkinliklere geçilir.

### 2. Laboratuvar öncesi etkinlikler

Öğrencilere konu ve imkanlar dahilinde neyi araştırmak istedikleri sorulur. Bu aşamada öğrencilerden araştırma sorularını yazmaları istenir. Araştırma soruları araştırılabilirlik açısından bütün sınıf tarafından eleştirilir. Öğrenciler araştırma soruları sınıf üyeleri tarafından uygun bulunursa hipotez, bağımlı değişken, bağımsız değişken, sabit değişken, deney malzemeleri ve veri toplama süreclerini rapor ederek deneylerine başlarlar.

### 3. Laboratuvar etkinliğine katılma

Bu aşamada öğrenciler gruplar halinde bir önceki aşamada belirlemiş oldukları deney sürecini takip ederek araştırma sorularına cevap ararlar. Öğretmen öğrencilerden bireysel olarak her türlü bulguyu raporlarına kayıt etmelerini ister.

### 4. Müzakere Fazı-1

Öğrenciler deneylerini tamamlayıp verilerini topladıktan sonra bireysel olarak iddia ve kanıtlarını oluştururlar. Bu iddia ve kanıtlar argümanlarını oluşturur ve bu argümanlar başlangıçta sormuş oldukları araştırma sorularına cevap niteliğindedir. Bu aşamada öğretmen öğrenciler ile etkileşime girerek onlara sorular sorabilir ve motive edebilir.

### 5. Müzakere Fazı-2

Öğrenciler bireysel olarak argümanlarını oluşturduktan sonra grup üyeleri ile grup içi tartışmasına girerek ortak bir grup argümanı oluştururlar. Öğrenciler tartışmaya girmedikleri zaman öğretmen grup ile etkileşime girerek deney ile ilgili veya teorik sorular sorabilir. Benzer şekilde öğretmen grup üyelerinin farklı bakış açılarına sahip olması için onlara rehberlik edebilir. Bu şekilde grup içi tartışmalar daha fazla yapılabilir.

**6. Müzakere Fazı-3**

Gruplar argümanlarını oluşturduktan sonra, bu aşamada gruplar sıra ile araştırma sorularını, deney süreçlerini, grup argümanlarını sınıfın geri kalanına sunarlar. Bu aşamada grup argümanı bütün sınıf tarafından tartışılır ve tüm sınıf argümantasyon sürecine dahil olur. Öğretmen bu aşamada tartışmanın verimli bir şekilde ilerlemesi için sorular sorar, önerilerde bulunur fakat kendisini dersin merkezinde konumlandırmaz. Süreç boyunca öğrenciler aktif ve özgür bir şekilde hareket ederler.

**7. Müzakere Fazı-4**

Sınıf tartışması tamamlandıktan sonra öğrenciler bu derste neler öğrendiklerini öğrenci raporunun yansıtmalar kısmına rapor eder. Bu şekilde öğrenci kendi öğrenmesinin farkına varır.

**8. Değerlendirme**

Değerlendirme süreç boyunca yapılabilir. Örneğin dersin başında öğrencilerin ön bilgisi açığa çıkartılır. Tartışmalar süresince öğrencilerin ne konularda eksikleri olduğu ortaya konur ve buna bağlı olarak öğretmen sürece dahil olabilir. Ayrıca öğrencilerin süreç boyunca yazdıkları SWH öğrenci raporu, öğrencilerin değerlendirilmesi açısından önemli bir veri kaynağıdır. Öğrenme süreci tamamlandıktan sonra öğrenciler alternatif yaklaşımlarla değerlendirilmeye devam edebilirler. Örneğin; öğrenciler bir alt sınıfta bulunan öğrencilere bu hafta öğrendikleri konu ile ilgili mektup yazabilirler.

## C. SAMPLE LESSON PLAN ABOUT CORE IDEAS INCLUDING EXPERIMENT

Konu Kazanımı: Isı iletimi ile ilgili deneyler yapıp tartışarak maddeleri ısı iletkenliğine göre sınıflandırır.

**1. Öğrencilerin ön bilgilerinin ortaya çıkarılması**

Öğretmen derse girer ve dersten önce hazırladığı sunuyu açar. Sunuda: Hafta sonu Kızılcahamam'a kampa gideceğini fakat götürdüğü yemeği kamp ateşinde nasıl pişirmesi gerektiğini bilmediğini söyler ve öğrencilerden bu konuda yardım ister. Yemeğin hızlı pişmesi için çelik kap mı yoksa toprak kap mı kullanılması gerektiği öğrencilere sorulur. Benzer şekilde yemeğin karıştırılması sırasında elimizin yanmaması için kullanılacak olan çatalın demir mi, tahta mı yoksa plastik mi olması gerektiği öğrencilere sorulur. Bu şekilde öğrencilerin ısının iletim yoluyla ilgili ön bilgileri ortaya çıkartılır. Öğrencilerden beklenen cevap yemeğin hızlı pişmesi için demir kap kullanılması ve elimizin yanmaması için tahta veya plastik çatal kullanılmasıdır. Öğrenci önbilgisi alındıktan sonra laboratuvar öncesi etkinliklere geçilir.

**2. Laboratuvar öncesi etkinlikler**

Öğrenciler maddelerin ısı iletimi ile ilgili araştırmak istedikleri soruyu grup olarak tartışırlar ve her grup araştırma sorusunu tahtaya yazar. Öğretmen araştırma sorusunu açık hale getirmek için öğrencilere neyi nasıl araştırmak istediklerini sorar. Bu durum uygulanacak olan deneyin grup tarafından anlaşılmasına yardımcı olur. Öğrenciler yapacakları deneyde bağımlı ve bağımsız değişkenlere karar verirler. Neleri sabit tutacaklarını rapor ederek bu değişkenlerin kontrol değişkeni olduğunu ifade ederler.

Fakat dersin kazanımı ile ilgili olmayan araştırma soruları ders kapsamından çıkartılır ve öğrencilerin programdan uzaklaşması engellenir.

Benzer şekilde sınıfta yapılan deneyle cevap verilemeyecek sorular araştırma sorusu olarak kullanılmaz. Örneğin bu konu ile ilgili "Neden demir kaşık tahta kaşıktan daha iyi bir iletkendir?” sorusu araştırma sorusu olarak kullanılamaz. Çünkü öğrenciler deney yaparak demir atomlarının tahtayı oluşturan atomlardan daha sık olduğunu gözlemleyemez. Ancak verilere bakarak böyle bir çıkarım yapabilirler. Dolayısı ile ilgili araştırma sorusu: “Tahta kaşık mı yoksa demir kaşık mı daha iyi iletkendir?” şeklinde yada “Maddelerin cinsi ile ısı iletkenlikleri arasında ilişki var mıdır?” sorusu ile değiştirilebilir.

**3. Laboratuvar etkinliğine katılma**

Araştırma sorusunu belirleyen ve değişkenlerine karar veren grup deneyi nasıl yapacağına karar vererek deneye başlar. Isının iletimi ile ilgili örnek olarak tasarlanan deney şu şekildedir:

Araştırma sorusu: Farklı maddeler ısıyı farklı şekillerde mi iletir?

Araştırma sorusu, hipotez ve değişkenler SWH raporunun başlangıç düşünceleri kısmına yazılır.

Hipotez: Maddenin türüne bağlı olarak ısı iletkenliği değişir.

Bağımsız değişken: (Isıyı ileten) madde cinsi

Bağımlı değişken: (Farklı maddelerin) ısı iletim hızı

Sabit (Kontrol) değişkeni: Madde üzerine konan yağın miktarı, boncuğun büyüklüğü, boncukların yerden yüksekliği

**Deney malzemeleri:**

1. Metal kaşık
2. Tahta kaşık
3. Cam çubuk
4. Plastik çubuk
5. Beherglas
6. Sıcak su
7. Termometre

8. Katı yağ
9. Boncuk

Öğrenciler deney prosedürünü kendileri belirler. Bu süreçte kendilerinden beklenen beherglasın içini sıcak su ile doldurmaları ve metal, tahta, cam ve plastikten yapılan maddeleri beherin içine koymalarıdır (Burada öğretmen yardımcı olabilir ayrıca suyu dökmemeleri konusunda öğrenciler uyarılır. Sıcak su tehlikeli olduğu için öğretmen bir termos getirip içindeki sıcak suyu behere koyabilir, öğrenci bu sürece katılmayabilir.). Daha sonra her bir maddenin ucuna katı yağ konur ve ısının etkisi ile yağın sıvılaşmasının gözlemlenmesinin kolay olması için yağların üzerine boncuklar konur. (Katı yağ ve boncuk kullanımı grubun kendi isteği ile olacaktır. Grup üyeleri neden boncuk ve yağ kullanacaklarını tartışabilirler.) Öğrenciler prosedür kısmında bu durumu akıllarına getiremeyebilecekleri için bir önceki fazda öğretmen deneyin nasıl tasarlanacağı konusunda öğrencilere yardım edebilir (katı yağ kullanımı, boncuk kullanımı ve bu maddelerin kullanım amacı gibi). Beher glasın içine termometre konarak suyun sıcaklığı ölçülür. Isı kaybını önlemek için beher glass alüminyum folyo ile kapatılır. Öğrenciler deneyi yaparken izledikleri prosedürü SWH raporunun test kısmına yazarlar.

**4. Müzakere Fazı-1**

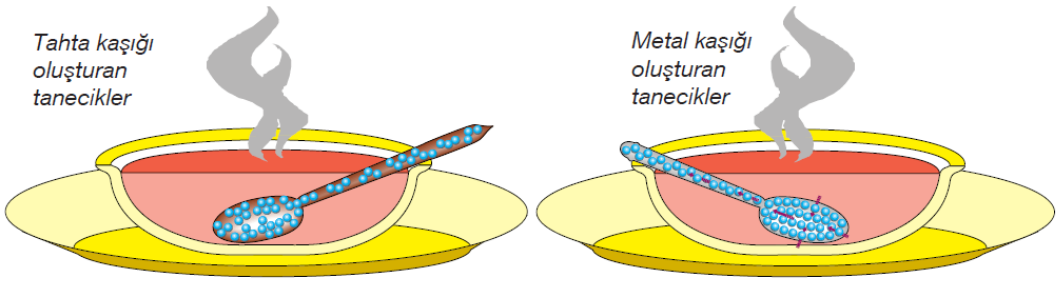
Öğrenciler deneyi yaparken gözlemlerini SWH öğrenci raporuna kayıt ederler. SWH öğrenci raporunun gözlemler ve bulgular (neleri buldum ve gözlemledim?) bölümü bireysel olarak cevaplanabilir. Örnek olarak verilen deneyde gözlemler süresince öğrenciler önce demirin üzerindeki katı yağın eridiğini ve bu maddenin üzerindeki boncuğun yere düştüğünü gözlemler ve bunu rapor ederler. Bu gözlemden hareketle öğrenciler bir iddiada bulunur ve bunu raporlarına kayıt ederler. İddia gözlemlere bağlı olarak ortaya atılır ve şu şekilde bir iddiada bulunulabilir (Deney sonunda ulaşılan sonuç): Demir kaşık diğerlerine göre ısıyı daha iyi iletir ya da maddeler ısıyı iletkenliklerine göre sınıflandırılabilir. Çünkü gözlemlere göre sıcak suda ısı en hızlı demir kaşıkta iletilmiş ve bunun neticesinde demir kaşığın üzerinde bulunan katı yağ en hızlı erimiş ve burada yer alan boncuk ilk düşmüştür. Diğer kaşıklar ile ilgili iddialar da benzer şekilde gözlemlere bağlı olarak verilebilir. Gözlem sonucu elde edilen bu veriler iddiaya kanıt oluşturmada kullanılabilir ve SWH öğrenci raporunun 5. bölümü olan kanıtlar kısmı bu şekilde doldurulur. Müzakere fazı-1 öğrenciler tarafından bireysel olarak yapılır.

**5. Müzakere Fazı-2**

Daha sonra her grup kendi içerisinde gözlemlerini birbiri ile paylaşır ve grubun tamamının ortak görüşü olan bir iddia ortaya konur. Örneğin; grup iddiası her madde ısıyı aynı şekilde iletmez olabilir, buna kanıt olarakta sıcaklık değişimine bağlı olarak yağın farklı maddeler üzerinde farklı hızlarda erimesi gösterilebilir. Benzer şekilde maddelerin ısı iletkenliği birbirleri ile karşılaştırılabilir (demir en hızlı iletti, plastik en yavaş iletti gibi).

**6. Müzakere Fazı-3**

Gruplar sunumlar yaparak neyi nasıl araştırdıklarını ve ne sonuca vardıklarını diğer gruplara anlatır ve diğer grupları ikna etmeye çalışır. Diğer gruplar ve öğretmen sordukları sorular ile sunum yapan grubu sınar ve gruplar arası tartışma ortamı oluşur. Bu aşamada örneğin sunum yapan gruba neden demirdeki yağın en hızlı eridiği öğretmen tarafından sorulabilir. Demirin ısıyı neden diğerlerinden daha iyi ilettiği sorusu da sınıfta tartışılabilir. Öğretmen yönlendirme yaparak öğrencilerin atomlar üzerinde düşünmesini sağlayabilir ve bu durumda demir atomlarının taneciklerinin daha düzenli ve sık olduğu sonucuna varılabilir. Buradan hareketle maddeler ısıyı çok ileten ve ısıyı az ileten maddeler olarak sınıflandırılır ve konu kazanımına ulaşılmış olur. Öğretmen burada öğrencilere atomları sık ve düzenli olduğu için ısıyı daha fazla ileten demir kaşık gibi maddelere ısıyı ileten madde, tahta kaşık gibi atomları düzensiz ve sık olmadığı için ısıyı daha az ileten maddelere de ısıyı yalıtan madde dendiğini söyleyebilir. Konunun somut hale getirilmesi için öğretmen aşağıdaki görseli kullanabilir. Öğrenciler SWH raporunda yer alan karşılaştırmalar kısmına diğer arkadaşlarının farklı düşüncelerini ve öğretmenin verdiği bilgileri yazarlar.




**7. Müzakere Fazı-4**

Bu aşamada konu ile ilgili olarak öğrenciler ders kitabından taneciklerin çarpışması ile ısının iletilmesi konusunu okuyarak konu ile ilgili daha fazla bilgi sahibi olabilirler (Örneğin; iletim yoluyla ısının iletimi sadece katı cisimlerde görülür) Benzer şekilde öğrenciler öğretmenin önerdiği EBA (http://www.eba.gov.tr/gorsel?icerikid=276694421e5ab23ad4ce69cc28d3bf09138a6 cfbbe002) gibi kaynakları kullanarak konu hakkında daha fazla bilgi sahibi olabilir. Okumalar ve videodan öğrenilenler de SWH raporunun karşılaştırmalar ve okumalar kısmına yazılabilir.

Okumalardan daha sonra öğrencinin kendi öğrenmesine yönelik yansımaları SWH raporuna yazması beklenir. Burada öğrenci düşüncelerim nasıl değişti sorusuna cevap arar. Yansımalar kısmının başarılı olması için öğretmen dersin başında gösterdiği ppt sunusunu tekrar açar ve kampta yemeği hızlı ısıtmak için hangi kabı kullanması gerektiğini ve yemek pişerken elinin yanmaması için hangi çatalın kullanılması gerektiğini nedenleri ile birlikte tekrar sorar. Öğrencilerin dersin başında bu soruya verdikleri cevap ile dersin sonunda bu soruya verdikleri cevabı ve aradaki farkları rapor etmeleri öğrencinin yansıma yaptığını gösterir. Öğrenci ayrıca süreç boyunca SWH raporu üzerinde yaptığı değişiklikleri ve neden bu değişiklikleri yaptığını da yansımalar kısmına rapor edebilir. Örneğin; bireysel düşünce ile ortaya atılan iddia grup içi tartışma sonucu SWH raporunda değişmişse öğrenci bu değişimi ve bu değişimin nedenini yansımalar kısmına rapor edebilir.

## 8. Değerlendirme

Değerlendirme hem süreç içinde hemde süreç sonunda yapılır. Dersin başında kullanılan ppt öğrencinin ilgili kazanıma ilişkin ön bilgisini yoklar. Dersin sonunda kullanılan ppt ise konu kazanımının süreç sonunda öğrenilip öğrenilmediğini değerlendirir. Süreç içerisinde ayrıca öğretmenin araştırma sorusu ve deney tasarımı ile ilgili sorular sorması süreç içi değerlendirmeye örnektir. Benzer şekilde müzakere-4'te sorulan neden demir kaşık ısıyı daha iyi iletir sorusu konu kazanımını ölçmeye yöneliktir. Son olarak öğrencilerin doldurdukları SWH raporu ders sonunda toplanır bu raporlar süreç sonunda yapılan değerlendirmeye örnek olabilir.

**Kaynak:**

Hand, B., Wallace, C. W., & Yang, Eun-Mi. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh grade science: Quantitative and qualitative aspects. International Journal of Science Education, 26(2), 131-149.

MEB. (2013). İlköğretim Fen ve Teknoloji Dersi (6, 7, 8. Sınıflar) Öğretim Programı. MEB, Ankara.

## D. SAMPLE LESSON PLAN ABOUT CORE IDEAS NOT INCLUDING EXPERIMENT

Kazanım: Yakıtları, katı, sıvı ve gaz yakıtlar olarak sınıflandırıp yaygın şekilde kullanılan yakıtlara örnekler verir.

Fosil yakıtların sınırlı olduğu ve yenilenemez enerji kaynaklarından biri olduğu belirtilir ve yenilenebilir enerji kaynaklarının önemi örnekler verilerek vurgulanır.

Farklı enerji kaynaklarını olumlu ve olumsuz özelliklerine göre tartışarak birbiri ile kıyaslar.

**1. Öğrencilerin ön bilgilerinin ortaya çıkartılması**

Öğretmen derse girer ve öğrencilere yolda giden bir araba, yanan bir lamba ve koşan bir insan resmi gösterir.



Yolda giden araç Yanan bir lamba Koşan insan

Öğretmen bu 3 resmin ortak özelliğini öğrencilere sorar. Öğrenciler tahminlerini yapar ve burada bir hareket olduğu veya iş yapıldığı gibi çıkarımlarda bulunurlar. Öğretmen öğrencilere bu eylemlerin gerçekleşmesi için neye ihtiyaç olduğunu öğrencilere sorar. Öğrencilerin enerjiye ihtiyaç olduğunu söylemeleri beklenir. Daha sonra ilgili resimlerde kullanılan enerjinin kaynağı ne olabilir şeklinde bir başka soru öğrencilere sorulur. Öğrenciler cevaplarını verdikten sonra farklı enerji kaynakları öğrencilerle paylaşılır. Bu enerji kaynakları: Taş kömürü, petrol, doğalgaz, güneş, rüzgar, yer altı ısısı (jeotermal kaynaklar), biyoyakıt (yaşayan veya yeni ölmüş bitkiler) ve odundur.

Daha sonra öğretmen sunusunu açarak öğrencilere bugünkü derste ‘Hangi enerji kaynağını kullanmalıyız?’ ve ‘Hangi enerji kaynağını kullanmamalıyız?’ konularını tartışacaklarını söyler.

**2. Etkinlik öncesi yapılanlar**

Öğrenciler hangi kaynağı kullanmamız gerektiği ile ilgili başlangıç düşüncelerini SWH öğrenci raporunun başlangıç düşünceleri kısmına nedenleri ile birlikte yazar. Araştırma sorusu da ‘Hangi enerji kaynağını kullanmalıyız?’ şeklinde başlangıç düşünceleri kısmına yazılır.

**3. Etkinliğe katılma**

Daha sonra öğretmen öğrencilere konu hakkında bilgi sahibi olmaları ve kendilerine sunulan verileri kanıt olarak kullanmaları ve analiz etmeleri için öğrencilere daha önce hazırladığı enerji kaynağı veri kağıdını verir. Bu kağıtta ilgili enerji kaynaklarının tanımları, avantajları ve dezavantajları yer almaktadır. Enerji kaynağı veri kağıdı aşağıda verilmiştir:

| **Enerji Kaynağı** | **Tanımı** | **Avantajları** | **Dezavantajları** |
|---|---|---|---|
| **Odun** | Yakıt olarak kullanılan odun ağaçlardan elde edilir. | Çok ekonomiktir<br><br>Kolay elde edilir. | Çabuk söner<br><br>Elde edilen ısı enerjisinin %75 i kullanılmadan duman ve küle dönüşür.<br><br>Havayı kirletir.<br><br>Küresel ısınmaya sebep olur<br>Asit yağmurlarına sebep olur |
| **Kömür** | Taş kömürü yerin altından çıkartılır.<br><br>Fosil yakıtlardır yani | Ekonomiktir<br><br>Kolay elde edilir. | Havayı kirletir<br><br>Küresel ısınmaya sebep |

| | | | |
|---|---|---|---|
| | günümüzden çok önce yaşamış bitki ve hayvanların yer kabuğunun derinliklerinde çürüyüp farklılaşmasıyla oluşmuştur. | | olur<br><br>Asit yağmurlarına sebep olur<br><br>Taşınması ve dağıtılması masraflıdır. |
| **Benzin/ Petrol** | Petrol, karalardan ve denizlerden sondaj makineleri ile çekilen sıvı fosil yakıtlardır.<br><br>Benzin, gaz yağı, mazot ve yağ yakıtlar petrolün ayrıştırılması ile ortaya çıkan enerji kaynaklarıdır. | Dağıtım sistemi gelişmiştir.<br><br>Kolay elde edilir | Taşınması ve dağıtılması pahalıdır.<br><br>Havayı kirletir.<br><br>Küresel ısınmaya sebep olur.<br><br>Talebe bağlı olarak fiyatları çok yüksek olabilir. |
| **Doğalgaz** | Petrol gibi yer kabuğunun kayaçları altında bulunur fakat gaz halindedir. | Dağıtım sistemi gelişmiştir.<br><br>Kolay elde edilir<br><br>Yakıldığında petrol, odun ve kömüre göre daha az karbondioksit üretir | Taşınması ve dağıtılması pahalıdır.<br><br>Havayı az kirletir.<br><br>Talebe bağlı olarak fiyatları çok yüksek olabilir. |
| **Güneş** | Güneş enerjisi güneş panelleri sayesinde kullanılabilir hale gelir.<br><br>Güneş panelleri gelen enerjiyi soğurarak diğer enerji çeşitlerine çevirir. | Güneş olduğu sürece bedavadır.<br><br>Evleri ısıtır<br>Evlerde bulunan sıcak su ihtiyacını karşılar.<br><br>Güneş fırını olarak gıdaları ısıtır. | Güneş her zaman bulunmayabilir.<br><br>Güneş almayan bölgeler bu kaynaktan faydalanamaz. |
| **Rüzgar** | Rüzgar enerjisi rüzgar türbini sayesinde kullanılabilir hale | Eğer hava rüzgarlıysa bedavadır. | Rüzgar her zaman bulunmayabilir. |

| | | | |
|---|---|---|---|
| | gelir.<br><br>Rüzgar türbinleri rüzgarın sahip olduğu hareket enerjisini elektrik enerjisine çevirebilir. | Küresel ısınmaya ve asit yağmuruna neden olmaz. | Deniz kenarına veya yüksek yerlere kurulduğu için her yere kurulamaz.<br><br>Çok fazla alan gerekir.<br><br>Bakımı ve onarımı pahalıdır.<br><br>Rüzgar ve fırtınalar türbinlere zarar verebilir. Çevrede yaşayan kuşların ölümüne sebep olabilir. |
| **Yeraltı ısısı (Jeotermal enerji kaynağı)** | Yerkürenin derinliklerinde bulunan ısıyı yüzeye çıkartıp ondan faydalanmamızı sağlar. | İşletme masrafları düşüktür<br><br>Evleri ısıtmak için kullanılabilir.<br><br>Jeotermal enerji kaynağı olan ısı eğer yüzeye yakınsa bu enerjiden elektrik üretilebilir. | Enerji tesisi kurmak masraflıdır. |
| **Biyoyakıt (yaşayan veya yeni ölmüş bitkiler)** | Bitkisel ve hayvansal atıkların yakılmasıyla enerji üretmemizi sağlar. | Düşük maliyetlidir.<br><br>Kırsal ve uzak alanlarda kullanılabilir.<br><br>Atık işletim sistemleri ile zararlı maddeleri ve çöpleri arıtmamıza yardımcı olur. | Küresel ısınmaya neden olur.<br><br>Yandığında sağlığa zararlı kurşun gibi maddeler havaya karışır. |

| | | | |
|---|---|---|---|
| | | Isınma amaçlı kullanılabilir. | |

SWH raporu 2. sorusu olan test, veri kaynakları olarak değiştiği için öğrenciler bu soruya hangi enerji kaynağını kullanmamız gerektiği sorusuna cevap bulmak için nasıl bir yol izlediklerini yazarlar. Örneğin; öğrenciler veri kağıdında yer alan avantajları ve dezavantajları kriter olarak kabul edip, bu kriterlere uygun olarak veriyi analiz edeceklerini SWH raporunun test kısmına yazabilirler.

**4. Müzakere Fazı-1**

Öğrenciler bireysel olarak veri kağıdından hangi enerji kaynağının diğerlerine göre daha fazla avantajı olduğu ile ilgili verileri rapor eder. Bunu SWH öğrenci raporunun gözlemler ve bulgular kısmına yazar. Daha sonra, en iyi enerji kaynağı bence budur diyerek bir iddiada bulunur, bu iddialarını ve gerekçelerini SWH öğrenci raporunun iddia ve kanıtlar kısmına yazar. Örneğin; öğrenci en iyi enerji kaynağının doğalgaz olduğunu iddia eder çünkü doğalgazın dağıtım sistemi gelişmiştir, kolay elde edilir ve diğer yakıtlara göre yakıldığında havayı daha az kirletir.

**5. Müzakere Fazı-2**

Öğrenciler bireysel iddialarını dile getirdikten sonra grup olarak hangi enerji kaynağının kullanılması ile ilgili tartışmaya girerler. Örneğin; doğalgaz kullanılmalı diyen öğrenciye karşı argüman olarak grup içerisinde bir başka öğrenci biyoyakıtın kullanılmasının daha iyi olduğunu dile getirebilir ve gerekçesini ise şu şekilde açıklar: “Biyoyakıt doğalgazdan daha ucuzdur ve devletin enerji kaynağını götüremediği yerlerde bile biyoyakıt kullanılabilir. Her yerde doğalgaza erişim yoktur. Ayrıca atıkların temizlenmesinde de biyoyakıtlar işe yarar.” Bu düşünceyi çürütmek için ise doğalgazı savunan öğrenci biyoyakıtların çevreye zarar verdiğini ve bu yakıtların hem küresel ısınmaya neden olduğunu hemde insan sağlığına zarar verdiğini söyler. Tartışma sonunda bütün grubu temsil eden bir sonuca ulaşılır. Öğrenciler kendilerinden farklı düşünen arkadaşlarının düşüncelerini SWH öğrenci raporunun karşılaştırmalar kısmına yazar.

**6. Müzakere Fazı-3**

Daha sonra öğretmen gruplar arası tartışmayı başlatır. Her grup kendi düşüncesinin doğruluğunu nedenleri ile birlikte açıklamaya çalışır. Öğretmen herhangi bir yönlendirmede bulunmaz. Tartışmaya katılmayan grupları sürece dahil etmeye çalışır. Tartışmanın durduğu yerlerde konuya farklı açılardan yaklaşır. Örneğin; güneş panelinin olumsuz yanlarının tartışılmadığını görürse güneş almayan bölgelerde bu enerjinin yeteri kadar kullanılamadığını söyler ve bu durumla ilgili öğrencilerin fikirlerini sorar.

Tartışma tamamlandıktan sonra öğretmen ders kitabında yer alan yakıtlar ve yenilenebilir enerji kaynakları konusunu öğrencilerin okumasını ister. Öğretmen öğrencilere: "Ders kitabındaki okumalarda işlediğimiz bilgilerden farklı olarak bir ifadeyle karşılaştınız mı?" diye sorar. Öğretmenin buradaki amacı öğrencilerin enerji kaynaklarını yenilenebilir ve yenilenemez enerji kaynakları olarak sınıflandırmasıdır.

Bu sınıflandırmadan sonra ise ders kitabında verildiği şekli ile öğrencilerin yakıtları katı, sıvı ve gaz yakıtlar diye sınıflandırması beklenir. Bu yakıtlara örnekler verilir. Örneğin; kömür katı, benzin sıvı, doğalgaz ise gaz yakıttır. Bu bilgiler ders kazanımı olduğu için öğretmen kendisi ppt sunusunu tekrar açar ve öğrencilere yenilenebilir ve yenilenemez enerji kaynaklarını gösterir. Yenilenemez enerji kaynaklarının fosil yakıtlar olduğu vurgusunu yapar ve fosil yakıtların nasıl oluştuğunu öğrencilere anlatır. Ayrıca, yakıtları sınıflandırır. Öğrenciler öğretmenden öğrendikleri yeni bilgileri ve kitaptan edindikleri bilgileri SWH raporunun okumalar kısmına rapor eder.

**7. Müzakere Fazı-4**

Öğrenciler bu dersin başında neleri biliyorlardı ve dersin sonunda neleri öğrendiler, bu bilgileri SWH yansımalar kısmına yazarlar. Bu bilgileri yaparken başlangıç düşüncelerindeki tahminleri ve gerekçeleri ile dersin sonundaki düşünceleri arasındaki fark yansımalar kısmına dahil edilebilir.

## 8. Değerlendirme

Dersin başında gösterilen resimler öğrencilerin günlük hayatta yapılan etkinlikleri enerji konusu ile bağlayıp bağlayamadıklarını öğrenmek amacıyla yapılacaktır. Bu görseller ayrıca konuyu enerji kaynakları ile ilişkilendirebilmek için kullanılacaktır. Öğrencilerin grup içerisinde ve gruplar arası yaptıkları tartışmalar öğretmeni süreç boyunca öğrencilerin ne düşündüğü ile ilgili bilgilendirir. Öğretmen tartışmadan sonra ders kitabının ilgili bölümlerini okutarak öğrencilerin bugünün kazanımı olan "Yakıtları, katı, sıvı ve gaz yakıtlar olarak sınıflandırıp yaygın şekilde kullanılan yakıtlara örnekler verir. " ifadesi ile ilgili çıkarım yapmalarını bekler ve bununla ilgili sorular sorar. Öğretmen, ders sonunda öğrencilerin hazırlamış olduğu SWH raporlarını toplayarak süreç boyunca öğrencilerin neler öğrendiklerini değerlendirir.

**Kaynakça:**

Hand, B., Wallace, C. W., & Yang, Eun-Mi. (2004). Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh grade science: Quantitative and qualitative aspects. International Journal of Science Education, 26(2), 131-149.

MEB. (2013). İlköğretim Fen ve Teknoloji Dersi (6, 7, 8. Sınıflar) Öğretim Programı. MEB, Ankara.

## E. MATTER AND HEAT CONTENT KNOWLEDGE TEST (MHCKT)

**1.** Eğer metal bir kaşık ile tahta bir kaşığı kaynamış su dolu bir kaba boşaltırsak birisi diğerinden daha fazla ısınır. Bu durumun sebebi nedir?
A) Metal ısıyı tahtadan daha iyi iletir.
B) Tahta ısıyı metalden daha iyi iletir.
C) Isı metal tarafından çekildiği için metaller ısınır.
D) Tahta metal kadar güçlü değildir.

**2.** Ocağın üzerinde kaynatılmakta olan bir cezvenin hangi maddelerden yapılmış olmasını beklersiniz?
A) Cezvenin tabanı ve elle tutulan kısmı ısıyı iyi ileten maddelerle kaplanmalıdır.
B) Cezvenin tabanı ve elle tutulan kısmı ısıyı iletmeyen seramik gibi yalıtkanlarla kaplanmalıdır.
C) Cezvenin tabanı ısıyı ileten maddelerle, elle tutulan kısmı ise ısıyı iletmeyen maddelerle kaplanmalıır.
D) Cezvenin tabanı ısıyı iletmeyen maddelerle, elle tutulan kısmı ise ısıyı iyi ileten maddelerle kaplanmalıdır.

**3.** Bir öğrenci aynı uzunlukta ve şekilde mermer ve bakır parçasının uçlarından tutarak bu iki maddenin diğer uçlarını özdeş (birebir aynı) ısıtıcılar yardımı ile ısıtmaktadır. Bu öğrenci aşağıdaki araştırma sorularından hangisine cevap arıyor olabilir? **(SPS-Research Question)**
A) Isı iletimi kullanılan maddenin uzunluğuna göre değişir mi?
B) Isı iletimi kullanılan maddenin şekline göre değişir mi?
C) Isı iletimi üzerinden ısı geçen maddenin türüne göre değişir mi?
D) Isı iletimi kullanılan ısı kaynağının türüne göre değişir mi?

**4.** Aşağıdaki tabloda bazı ısıtma kapları ve bu kapların özellikleri verilmiştir:

| **Isıtma Kabının Türü** | **Isıtma Kabının Özellikleri** |
|---|---|
| Toprak Kap | Isıyı ağır ağır yayarak yemekleri uzun sürede pişirir. Toprağın yapısında bulunan zararlı kimyasal maddeler pişirme sırasında yemeklere karışabilir. Toprağın ısı iletkenliği havanın 25 katıdır. |
| Cam Kaplar | Cam pişirme kapları sağlık açısından sorun yaratmaz. Camın ısı iletkenliği havanın 42 katıdır. |
| Çelik Tencere | Çeliğin pişirme süresi uzundur. Çelik tencereler aşınırsa, tencere ısınırken çeliğin yüzeyinde bulunan metaller yemeğe karışıp sağlık sorununa sebep olabilir. Çeliğin ısı iletkenliği havanın 534 katıdır. |
| Alüminyum tencere | Yapılan araştırmalar alüminyumun kansere ve bazı sinir hastalıklarına neden olduğunu göstermektedir. Alüminyumun ısı iletkenliği havanın 8300 katıdır. |

Yukarıdaki bilgiler göz önüne alındığında yemeğinin yavaş pişmesini isteyen ve yemeğin sağlığına zarar vermesini istemeyen bir kişi hangi ısıtma kabını tercih etmelidir?

A) Toprak kabı seçmelidir çünkü toprak kap ısıyı çok fazla iletmez ve sağlığa zarar vermez.

B) Cam kap seçilmelidir çünkü cam kap ısıyı çok fazla iletmez ve sağlığa zarar vermez.

C) Çelik tencere kullanılmalıdır çünkü çelik tencere ısıyı çok fazla iletmez ve sağlığa zarar vermez.

D) Alüminyum tencere kullanılmalıdır çünkü alüminyum tencere ısıyı çok fazla iletmez ve sağlığa zarar vermez.

**5.** Binalarda ısı yalıtımı yapılırken bazı ölçütlere dikkat edilir. Isı yalıtımı yapılırken aşağıda yer alan ölçütlerden hangisine dikkat edilmez?

A) Kullanılan madde yüksek sıcaklıklarda özelliğini kaybetmemeli

B) Kullanılan madde dış etkenlere karşı sağlam olmalı

C) Kullanılan madde ısı akışını yavaşlatmalı

D) Kullanılan madde ekonomik olarak pahalı olmalı

**6.** "Kışın evlerinin soğuk olmasını istemeyen apartman sakinleri bir araya gelerek binalarının yalnızca dış cephesine ısı yalıtımı yaptırmak istemekte ve evlerinin kapı, zemin gibi diğer bölümlerine ısı yalıtımı yaptırmamaktadır. Isı yalıtımı yapan ustalara ev sahipleri ustaların hangi yalıtım malzemesini kullanacağını sorar. Ustalar dış yalıtım için köpük kullanacaklarını, eğer zemini de yalıtım malzemesi ile kaplatmak isterlerse zemin için taş yünü malzemesini kullanabileceklerini söylerler."

Yukarıdaki durumdan hareketle aşağıdaki sonuçlardan hangisine ulaşılamaz?

A) Evin zemininde de evin diğer yerleri gibi ısı akışı gerçekleşmektedir.

B) Isı yalıtımı yapılan evlerde kışları evlerin soğuk olması engellenir.

C) Isı yalıtımı yaparken kullanılacak malzemenin kolayca yanmaması gerekir.

D) Isı yalıtımı yaparken kullanılacak malzemeler yalıtım yapılacak yere göre farklılık gösterir.

**7.** Bir öğrenci soğuk kış günlerinde içtiği sıcak çayın soğumaması için bir termos yapmak istemektedir. Aynı öğrenci ışık ışınlarının parlak yüzeylerden geri yansıdığını ve parlak olmayan koyu renkli yüzeylerin ışık ışınlarını emerek yüzeyi ısıttığını bilmektedir. Ayrıca öğrenci iki madde arasında bulunan diğer maddenin de ısı akışına neden olduğunu bilmektedir. Bu bilgilerin yanısıra öğrencinin yapmış olduğu termosun özellikleri şu şekildedir:

- Termos iç içe geçmiş 2 adet cam bardaktan oluşmaktadır.
- Dışarıdaki cam bardak parlak alüminyum folyo ile kaplıdır.
- İçerideki cam bardak parlak alüminyum folyo ile kaplıdır.

Öğrenci sıcak çayı soğuk bir günde iç kısımda kalan bardağa koyar. Çayı koyduktan sonra termosun üzeri kağıt ile kapatılır. Bu öğrencinin çayını daha uzun süre sıcak tutabilmesi için ne yapması gerekmektedir?
A) Bardakların arasında kalan boşluğu soğuk su ile doldurmalıdır.
B) Bardakların arasında kalan boşluğu kağıt parçaları ile doldurmalıdır.
C) Dışarıdaki bardağı kaplayan alüminyum folyo siyah kartonla değiştirilmelidir.
D) İçeride yer alan cam bardak metal bardak ile yer değiştirmeli.

**8.** Çantasında yer alan bir kutu kolayı soğuk tutmak isteyen bir öğrenci kutu kolayı hangi madde ile kaplamalıdır, neden?
A) Alüminyum folyo ile kola kaplanmalıdır çünkü metaller ısıyı iyi iletirler.
B) Kutu kola kağıt havlu ile kaplanmalıdır çünkü kağıt nemi çeker.
C) Kutu kola balmumu kağıdı ile kaplanmalıdır çünkü balmumu nemi emer.
D) Kutu kola yün kazak ile kaplanmalıdır çünkü yün havayı hapseder.

**9.** Bir öğretmen öğrencilerinden yaz aylarında kullanılmak üzere bir ayakkabı tasarlamalarını istemiştir. Amaç yazın sıcak havada ayağımızı tasarlanan ayakkabı yardımıyla serin tutmaktır. Bir öğrenci grubunun tasarlayacakları ayakkabılarla ilgili düşünceleri aşağıdaki gibidir:
1. öğrenci: Malzeme olarak ayakkabının alt kısmı plastikten yapılabilir.
2. öğrenci: Kullanacağımız malzemenin içerisinde boşluklar olmalıdır.
3. öğrenci: Kullanacağımız malzeme koyu renkli olmalıdır.
4. öğrenci: Kullanılacak ayakkabı siyah asfalt zeminde kullanılmalıdır.
Yukarıda düşünceleri verilen öğrenciler ile ilgili aşağıdaki ifadelerden hangisi doğrudur?
A) 1. öğrenci yanlış düşünmektedir çünkü yazın ısınan zemin plastik tabandan ayağa ısı transferi ederek ayağı rahatsız eder.
B) 2. öğrenci doğru düşünmektedir çünkü malzemenin içerisindeki boşlukta hava birikir ve ayakkabının dışından içeri doğru ısı transferi hava ortamında fazla olmaz.
C) 3. öğrenci doğru düşünmektedir çünkü güneş ışınları koyu renkli maddelerden yansıyarak geri döner ve bu durum ayakkabının ısınmasını engeller.
D) 4. öğrenci doğru düşünmektedir çünkü güneş ışınları siyah asfalttan yansıyarak asfaltın soğumasına neden olur ve bu durumda ayakkabı ısınmaz.

**10.** Aynı özellikteki iki binanın dış cephelerine ısı yalıtımı yapılacaktır. Binalarda kullanılacak ısı yalıtım malzemelerinin havaya göre ısı iletkenlik katsayıları ve malzeme kalınlığı aşağıdaki tabloda verilmiştir:

| Binalar | Dış Cephe malzemesi | Malzemenin havaya göre ısı iletme katsayısı | Binada Kullanılan Malzeme kalınlığı |
|---|---|---|---|
| 1. Bina | Cam yünü | 40 katı | 1 cm. |
| 2. Bina | Polistren köpük | 10 katı | 4 cm. |

Kullanılan dış cephe malzemeleri ve malzeme kalınlığı göz önüne alındığında aşağıdaki çıkarımlardan hangisi yapılamaz?

A) Cam yünü polistren köpükten daha iyi bir ısı yalıtıcısıdır.
B) Dış cephe mazlemesi olarak farklı maddeler kullanılabilir.
C) Binalara yapılan dış cephe yalıtımı farklı kalınlıkta olabilir.
D) Cam yünü polistren köpüğe göre ısıyı daha iyi iletir.

**11.** Aşağıdaki tabloda Türkiye'de 3 farklı bölgede bulunan binalarla ilgili yalıtımlı bina oranı, yalıtımlı binaların kalınlığı ve o bölgede gerçekte olması gereken yalıtım kalınlığı verilmiştir.

| Bölge | Yalıtımlı Bina Oranı | Binaların Yalıtım Kalınlığı | Olması gereken yalıtım Kalınlığı |
|---|---|---|---|
| Marmara | %60 | 5 cm | 10 cm |
| İç Anadolu | %70 | 6 cm | 12 cm |
| Akdeniz | %15 | 4 cm | 8 cm |

Bu bilgiler göz önünde bulundurulduğunda aşağıdaki sonuçlardan hangisine ulaşılamaz?

A) Türkiye'de yalıtımlı bina oranı bölgeden bölgeye farklılık göstermektedir.
B) Akdeniz bölgesi diğer bölgelerden daha sıcak olduğu için yalıtımlı bina oranı daha azdır.
C) Türkiye'de yalıtımlı binaların yalıtım kalınlığı olması gerekenden daha incedir.
D) Olması gereken bina yalıtım kalınlığı ülkemizde bölgeden bölgeye farklılık göstermektedir.

**12.** Aşağıdaki ifadelerden hangisi ısı yalıtımının ülke ekonomisi ve kaynakların etkili kullanımı ile ilgili doğru bir ifade değildir?

A) Isı yalıtımı yapılan yerlerde kalın malzeme kullanmak ısı akışını keseceği için ince yalıtım malzemesi kullanılmamalıdır.
B) Isı yalıtımının yapılması ile fosil yakıt kaynakları daha az kullanılacağı için hava kirliliği azalacaktır.
C) Isı yalıtımı yapılan yerlerde daha az yakıt yakılacağı için ısı yalıtımı ekonomik masrafları azaltacaktır.
D) Sıcak bölgelerde ısı yalıtımı yaptırmaya gerek yoktur bu bölgelerde yapılan ısı yalıtımı gereksiz ekonomik masraf getirecektir.

**13.** Aşağıdaki bilgilerden yanlış olan seçeneği işaretleyiniz.
A) Fosil yakıtlar sınırsız enerji kaynaklarıdır.
B) Odun, kömür ve doğalgaz fosil yakıtlara örnek verilebilir.
C) Fosil yakıtlar yakıldığında bir enerji türü başka bir enerji türüne dönüşebilir.
D) Fosil yakıtlar katı, sıvı ve gaz yakıtlar olmak üzere üçe ayrılır.

**14.** Enerji kaynaklarını sınıflandırmak isteyen bir öğrenci enerji kaynaklarını 2 sınıfa ayırarak aşağıdaki tabloyu oluşturmuştur:

| 1. Sınıf | 2. Sınıf |
|---|---|
| Kömür | Rüzgâr |
| Benzin | Jeotermal |
| Doğal gaz | Güneş |

Yapılan bu sınıflandırma göz önünde bulundurulduğunda aşağıdaki sonuçlardan hangisine ulaşılabilir?
A) Enerji kaynakları yer altı ve yer üstü enerji kaynakları olarak sınıflandırılabilir.
B) Enerji kaynakları yenilenebilir ve yenilenemez enerji kaynakları olarak sınıflandırılabilir.
C) Enerji kaynakları katı, sıvı ve gaz enerji kaynakları olarak sınıflandırılabilir.
D) Enerji kaynakları denizden ve karadan elde edilen enerji kaynakları olarak sınıflandırılabilir.

**15.** Aşağıdaki tabloda farklı enerji kaynaklarının çevreye verdikleri zarar ve bu kaynakların kullanımı için gerekli ekonomik masraflar verilmiştir.

| Enerji Kaynağı | Enerji Kaynağının çevreye verdiği zarar | Ekonomik Masrafı |
|---|---|---|
| Güneş Paneli | Zararsız | Çok az |
| Rüzgar Gülü | Çok az | Orta |
| Petrol | Çok fazla | Fazla |
| Doğalgaz | Orta | Fazla |

Yukarıdaki bilgiler göz önüne alındığında aşağıdaki sonuçlardan hangisi <u>çıkartılamaz?</u>
A) Yenilenebilir enerji kaynaklarının maliyetleri ve çevreye verdikleri zarar benzerdir.
B) Çevreye en az zarar veren ve en ucuz enerji kaynağı güneş panelidir.
C) Yenilenemez enerji kaynakları yenilenebilir enerji kaynaklarından daha fazla çevreye zarar verir.
D) Yenilenemez enerji kaynakları yenilenebilir enerji kaynaklarından daha pahalıdır.

**16.** Fosil yakıtların yanısıra birde biyoyakıt adı verilen bir yakıt çeşidi vardır. Buna göre milyonlarca yıl önce ölmüş bitki ve hayvan kalıntıları fosil yakıtları oluştururken, yaşayan veya yeni ölmüş bitkiler biyoyakıtları oluşturmaktadır. Çevreye olan etkileri göz önüne alındığında biyoyakıt ve fosil yakıtlardan hangisinin kullanılması gerekmektedir?

A) Her ikisi de kullanılmalıdır çünkü bu yakıtlar ihtiyacımız olan enerjiyi karşılar.
B) Hiç birisi kullanılmamalıdır çünkü bu yakıtlar havayı kirletmektedir.
C)Biyoyakıtlar kullanılmalıdır çünkü biyoyakıtlar çevreyi kirletmeden önce havayı temizlemektedir.
D) Fosil yakıtlar kullanılmalıdır çünkü fosil yakıtlar havayı biyoyakıtlar kadar kirletmez.

**17.** Fosil yakıtların ısı amaçlı kullanımının insan ve çevreye karşı zararları vardır. Aşağıdakilerden hangisi fosil yakıtların insan ve çevreye verdiği zararlardan değildir?
A) Fosil yakıtların yakılmasıyla ortaya çıkan gazlar küresel ısınmaya sebep olur.
B) Fosil yakıtların yakılmasıyla ülkelerin ihtiyacı olan enerji karşılanmış olur.
C) Fosil yakıtların yakılmasıyla havadaki zararlı gazların oranı artar, bu gazların yağmurlarla yeryüzüne inmesi ile asit yağmurları oluşur.
D) Fosil yakıtların yakılması ile hava kirliliği oluşur ve bu kirlilik solunum yolu hastalıklarına sebep olabilir.

**18.** Son yıllarda Ankara'da kömür kullanımının artmasından dolayı şehrimizde hava kirliliği olması gereken sınırın üzerine çıkmıştır. Aşağıdaki ifadelerden hangisi hava kirliliğini azaltmaya yönelik önlemlerden birisi değildir?
A) Kömür kullanımı yerine çevreyi daha az kirleten yakıt tüketimi teşvik edilmelidir.
B) Kalitesiz kömür kullanımının önüne geçilmelidir.
C) Binalara ısı yalıtımı yapılarak evlerden ısı akışının önüne geçilmelidir.
D) İnsanlar toplu ulaşım yerine bireysel ulaşımı tercih etmelidir.

**19.** Aşağıdaki tablo kömür madeninin yer altından yer yüzüne çıkartılmasına kadar geçen sürede verdiği zararları göstermektedir.

| Çevreye verilen zarar | İnsanlara verilen zarar |
|---|---|
| Kömür yer yüzüne çıkartılırken yüzeyde bulunan ve yerin altında bulunan su kaynakları kirlenir. | Kömür çevredeki su kaynaklarını kirleterek içme suyumuza zarar verir ve bu durum insanların sağlığını tehdit eder. |
| Terk edilen madenlerde çıkan maden yangınları çevreye ve orada yaşayan bitki ve hayvanlara zarar verir. | Kömür çıkartan maden işçileri birçok solunum yolu hastalığına yakalanır çünkü kömür tozları işçilerin akciğerlerinde birikir. |

Yukarıdaki bilgiler göz önüne alındığında aşağıdaki sonuçlardan hangisi çıkartılamaz?

A) Kömür madeninin çıkartılması sırasında hem çevre hemde insanlar zarar görür.

B) Su kirliliği ve yangın tehlikesi kömür çıkartılırken çevreye verilen zararlara örnek olabilir.

C) Kömür madeninin çıkartılması insanların hastalanmasına neden olmaktadır.

D) Kömür madeni çıkartılırken çevreye verilen zarar insana verilen zararla ilgili değildir.

**20.** Aşağıda yer alan ifadelerde kömür ve doğalgaz ile ilgili bazı bilgiler yer almaktadır:

| |
|---|
| • Kömürün ve doğalgazın çıkartılması, taşınması ve dağıtılması sırasında para harcanır. |
| • Kömür yakıldığı zaman doğalgaza göre havayı daha fazla kirletir. |
| • Kömür ülkemizde bol miktarda bulunuyorken, doğalgaz ülkemizde çok az bulunmaktadır. İhtiyacımız olan doğalgaz ise diğer ülkelerden satın alınmaktadır. |

Bu bilgilere göre ülkemizin **ekonomisi** göz önünde bulundurulduğunda hangi yakıtın kullanılması gerektiği ile ilgili olarak aşağıdaki önerilerden hangisi doğrudur?

A) Kömür kullanılmalıdır çünkü kömür ülkemizde bol miktarda bulunmaktadır.

B) Doğalgaz kullanılmalıdır çünkü doğalgazı kullanmak için üreten ülkelerden satın almamız yeterlidir.

C) Kömür kullanılmalıdır çünkü kömürün çıkartılması ve dağıtılması sırasında para harcanmaz.

D) Doğalgaz kullanılmalıdır çünkü doğalgaz havayı çok fazla kirletmez.

## F. ELECTRICITY CONTENT KNOWLEDGE TEST (ECKT)

**1.** Bir elektrik devresinde yanmakta olan ampul vardır. Devreye iletken bir madde bağlandığında ampulün parlaklığı ile ilgili nasıl bir gözlem yaparız? (**Devreye bağlanan iletkenin direncini ihmal ediniz.**)
A) Değişmez
B) Artar
C) Azalır
D) Ampul söner

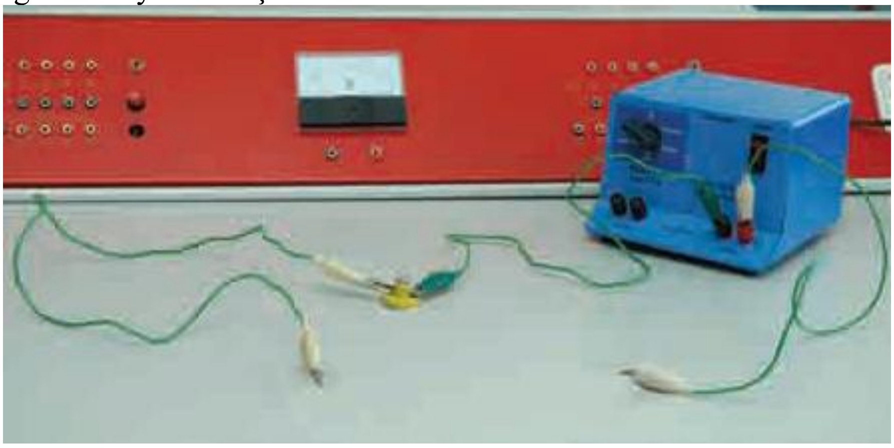
**2.** Bir elektrik devresinde yanmakta olan bir ampul vardır. Devreye yalıtkan bir madde bağlandığında ampulün parlaklığı ile ilgili nasıl bir gözlem yaparız?
A) Değişmez
B) Artar
C) Azalır
D) Ampul söner

**3.** Demir gibi metal cisimler elektrik akımını iletirken, silgi gibi maddeler elektrik akımını iletmezler. Fakat karbon gibi bazı maddeler de elektrik akımını kısmen ilettikleri için yarı iletken ismini alırlar. Bu ifadelerden aşağıdaki yargılardan hangisi çıkartılamaz?
A) Bazı maddeler elektrik akımını bazı maddelerden daha iyi iletir.
B) Bazı maddeler elektrik akımını iletmezler.
C) Maddeler elektrik akımını iletenler ve iletmeyenler olarak 2'ye ayrılırlar.
D) Elektrik akımını iletme özelliği maddeden maddeye değişebilir.

**4.** Bir öğrenci odasının kapısının demirden yapılmış kolunu her açtığında kendisini elektrik çarptığını gözlemlemektedir ve bu durum canını acıtmaktadır. Öğrenci bu soruna bir çözüm getirmek istemektedir. Aşağıdaki önerilerden hangisi bu sorunun çözümü olamaz?
A) Öğrenci kapının demir kolunu plastik kapı kolu ile değiştirmelidir.
B) Öğrenci metal bir tutacak ile kapıyı açmalıdır.
C) Öğrenci kapının demir kolunu bant ile kaplamalıdır.
D) Öğrenci plastik bir eldiven giyerek kapıyı açmalıdır.

**5.** Bir öğrencinin telefonunun şarj aleti bozulmuştur. Sorunu çözmek için öğrenci şarj aletinin kablosunun üzerindeki plastik kaplamayı soymuş ve şarj aletinin içinde yer alan metal telin kopuk olduğunu fark etmiştir. Öğrenci kopan teli bağladıktan sonra şarj aletinin çalıştığını gözlemlemiştir.
Bu bilgilere göre aşağıdaki önerilerden hangisi şarj aletinde yaşanan problemi çözmek için ortaya atılabilecek bir diğer öneridir?
A) Kopan metal teller yeni bir metal tel ile değiştirilebilir.
B) Şarj aletinin üzerindeki plastik kaplama yenisi ile değiştirilebilir.
C) Şarj aletinin fişi yenisi ile değiştirilebilir.
D) Şarj aletinin telefona bağlanan ucu yenisi ile değiştirilebilir.

**6.** Bir öğrenci aşağıdaki gibi kurulmuş olan bir elektrik devresinin boşta kalan kablolarına sırası ile plastik ataş, alüminyum folyo, çivi ve cam bağlamıştır. Devrede bulunan ampul alüminyum folyo ve çivi bağlandığında yanmış; plastik ataş ve cam bağlandığında ise yanmamıştır.



Bu bilgilere aşağıdaki sınıflamalardan hangisi doğrudur? **(SPS Classify)**

| | İletken | Yalıtkan |
|---|---|---|
| A) | Alüminyum Folyo | Çivi |
| B) | Cam | Alüminyum folyo |
| C) | Çivi | Plastik Ataş |
| D) | Plastik Alaş | Cam |

**7.** Bir öğrenci güç kaynağı, ampul ve elektrik devresi kablosu kullanarak bir deney hazırlamaktadır. Bu malzemeleri birbirine bağladığı zaman öğrenci ampulün yandığını gözlemlemiştir. Daha sonra, aynı öğrenci kabloları birbirinden ayırmış ve kabloları farklı sıvıların içine sokarak ampulün yanıp yanmadığını gözlemlemiştir. Buna göre kablolar tuzlu su ve sirkenin içine konduğunda ampul yanmış, şekerli su ve saf suyun içine konduğunda ampul yanmamıştır. Bu deneyi yapan öğrenci aşağıdaki araştırma sorularından hangisine cevap arıyor olabilir? **(SPS Research question)**

A) Elektrik iletiminde güç kaynağının görevi nedir?
B) Elektriğin iletimi için hangi devre elemanları gereklidir?
C) Elektriğin iletimi sıvılarda gerçekleşir mi?
D) Elektriğin iletimi sıvının cinsine bağlı mıdır?

**8.** Bir öğrenci iki farklı deney hazırlamaktadır. Her iki deneyde de ortak olarak aynı özellik ve sayıda güç kaynağı, ampul, elektrik devresi kablosu kullanılmıştır. 1. Deneyde devreye çivi bağlanmış ve ampulün yandığını gözlemlenmiş, 2. Deneyde devreye plastik ataş bağlanmış ve ampulün yanmadığı gözlemlenmiştir. Öğrencinin yapmış olduğu bu 2 deneye göre aşağıdakilerden hangisi kontrol değişkeni sayılamaz? **(SPS Identifying variables)**

A) Güç kaynağının sayısı
B) Ampulün çeşidi
C) Elektrik kablosunun türü
D) Birinci devreye bağlanan çivi

**9.** Bir öğrenci 2 ayrı deney yapmaktadır. Her iki deneyde de ortak olarak güç kaynağı, bir adet ampul ve elektrik devresi kablosu bulunmaktadır. Birinci deneyde bağlantı kablolarının arasına ek olarak metal çivi bağlayan öğrenci ampulün yandığını gözlemlemiştir. İkinci deneyde ise öğrenci elektrik devresi kablolarını birbirinden ayırmış ve kabloları tuzlu suya sokmuştur. Bu deney sonucunda da öğrenci ampulün yandığını gözlemlemiştir. Her iki deney birlikte düşünüldüğünde aşağıdaki sonuçlardan hangisine ulaşılabilir?

A) Katı cisimler elektriği sıvılardan daha iyi iletir.
B) Hem sıvılar hemde katılar elektriği iletebilir.
C) Elektrik iletimi sıvı cismin türüne bağlıdır.
D) Metal çivi iletken olduğu için devredeki ampul yanmıştır.

**10.** Metaller elektriği ileten maddelerdir ve her bir maddenin elektriği iletme katsayısı mega-siemens/m formulü ile hesaplanır. Elektrik iletme katsayısı fazla olan metaller elektriği elektrik iletme katsayısı düşük olan metallere göre daha fazla iletirler. Aşağıdaki tablo farklı metallerin elektrik iletkenlik katsayılarını göstermektedir.

| **Metal** | **Elektrik İletme katsayısı (mega-siemens/m)** |
|---|---|
| Gümüş | 63 |
| Alüminyum | 36 |
| Çinko | 16 |
| Demir | 9 |
| Çelik | 1 |

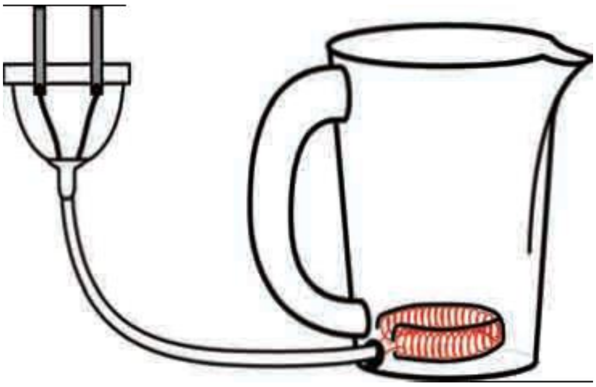
Yukarıdaki bilgiler birlikte düşünüldüğünde aşağıdaki **genellemelerden** hangisine ulaşılabilir?

A) Elektriği en iyi ileten metal gümüştür.
B) Alüminyum iyi bir iletkendir.
C) Metallerin iletkenlikleri metalin cinsine göre değişir.
D)Demirin elektrik iletkenliği çelikten daha iyidir.

**11.** Bir öğrenci maddeleri iletkenliklerine göre iletken ve yalıtkan olarak sınıflandırmak istemektedir. Bu düşüncesini gerçekleştirmek içinse bir deney tasarlamaktadır. Bu deneyde öğrenci bir güç kaynağının iki ucuna elektrik devresi kablosu bağlayıp, kabloların arasına da ampul bağlamıştır ve devre tamamlandığında (bu üç devre elemanı birbirine bağlandığında) ampul yanmaktadır. Öğrenci ilk olarak deney düzeneğine alüminyum folyo eklemiş ve ampulün yandığını gözlemlemiştir. İkinci durumda ise deney düzeneğinden alüminyum folyoyu çıkartıp düzeneğe demir anahtar eklemiş ve tekrar ampulün yandığını gözlemlemiştir. Bu iki durum göz önüne alındığında öğrencinin yapmış olduğu deneyin amacına ulaşıp ulaşmadığı konusunda ne söylenebilir? **(SPS Designing Experiment)**
A) Deney başarıya ulaşmamıştır çünkü her iki durumda da ampulün yanması deneyin sadece iletkenlerle yapıldığını göstermektedir.
B) Deney başarıya ulaşmamıştır çünkü kullanılan maddelerin iletken mi yoksa yalıtkan mı olduklarına dair elimizde bilgi yoktur.
C) Deney başarıya ulaşmıştır çünkü deneye bağlanan devre elemanları her iki durumda da ampulün yanmasını sağlamıştır.
D) Deney başarıya ulaşmıştır çünkü deneyde farklı devre elemanları kullanılmıştır.

**12.** Aşağıdaki durumlardan hangisi yapıldığında devrede bulunan ampulün parlaklığı <u>değişmez?</u>
A) Devrede bulunan güç kaynağı devreye ters bağlandığında ampulün parlaklığı değişmez.
B) Devrede bulunan güç kaynağının ucuna bir başka güç kaynağı eklendiğinde ampulün parlaklığı değişmez.
C) Devrede bulunan elektrik kablosunun boyu uzatıldığında ampulün parlaklığı değişmez.
D) Devrede bulunan ampulün yanına bir ampul daha eklendiğinde ampulün parlaklığı değişmez.

**13.** Tüm elektrik devre elemanları elektrik enerjisinin devrede iletilmesine karşı belli miktarda direnç uygularlar. Yani güç kaynağının, elektrik devresi kablosunun, ampulün de bir direnci vardır. Bazı devre elemanlarının direnci fazlayken bazı devre elemanlarının direnci çok azdır ve çok az direnç uygulayan devre elemanlarının direnci elektrik ile ilgili sorularda bazen ihmal edilir.
Yukarıdaki bilgilere göre; aşağıdaki ifadelerden hangisi doğrudur?
A) Elektrik devresi kablosu elektrik iletimine karşı direnç göstermez.
B) Elektrik devre elemanları elektrik iletimine karşı aynı direnci gösterirler.
C) Devrede yer alan güç kaynağının da bir direnci vardır.
D)Elektrik devresinde elektrik enerjisinin iletimine karşı en büyük direnci ampul gösterir.



**14.** Yukarıdaki şekil bir su ısıtıcını göstermektedir. Buna göre su ısıtıcısının fişinde yer alan tellerde küçük direnç, su ısıtıcısının zemininde yer alan tellerde ise büyük direnç vardır. Eğer su ısıtıcısının fişinin tellerinde büyük direnç, zemininde yer alan tellerde küçük direnç olsaydı aşağıdaki durumlardan hangisi **oluşmazdı?**
A) Fişin içindeki tel daha fazla ısınırdı.
B) Su ısıtıcısının zeminindeki tel daha fazla ısınırdı.
C) Fişin içerisinde taşınan elektrik enerjisi daha fazla zorlukla karşılaşırdı.
D) Isıtıcının zemininde taşınan elektrik enerjisi daha az zorlukla karşılaşırdı.

**15.** Bir öğrenci yapmış olduğu bir deneyde öncelikle 1adet elektrik devresi kablosu, 1 adet güç kaynağı ve 1 adet ampul kullanmaktadır. Öğrenci devreyi kurduğunda ampulün yandığını gözlemlemektedir. Daha sonra aynı öğrenci devreye 2 adet elektrik devresi kablosu daha ekler ve ampulün daha az parlak yandığını gözlemler. Bu deneyi yapan öğrencinin bağımsız değişkeni aşağıdakilerden hangisidir? **(sps identify variables)**
A) Ampulün sayısı
B) Elektrik devresi kablosunun uzunluğu
C) Güç kaynağının sayısı
D) Ampulün parlaklığı

**16.** Bir öğrenci ampulün parlaklığını etkileyen etmenleri araştırdığı bir deneyde öncelikle 1 adet ampul, 1 adet güç kaynağı ve 1 adet ince elektrik teli kullanmıştır. Bu durumda ampul yanmıştır. Daha sonra öğrenci deneyinde yer alan ince elektrik telini çıkartıp onun yerine aynı cinste fakat daha kalın elektrik teli kullanmıştır ve bu durumda ampulün daha parlak yandığını gözlemlemiştir. Bu deneyi yapan öğrencinin bağımlı değişkeni aşağıdakilerden hangisi olabilir? **(sps identify variables)**
A) Elektrik telinin kalınlığı
B) Elektrik telinin cinsi
C) Güç kaynağının sayısı
D) Ampulün parlaklığı

**17.** Bir öğrenci elektrik devresinde ampulün parlaklığını etkileyen etmenleri araştırmaktadır. Bunun için güç kaynağı, ampul ve elektrik devresi ile kurmuş olduğu devreye ilk önce gümüş tel eklemiş ve ampulün yanmaya devam ettiğini görmüştür. Aynı öğrenci daha sonra devreden gümüş teli çıkartıp devreye aynı uzunluk ve kalınlıkta demir tel eklemiş ve ampulün daha az parlak yandığını gözlemlemiştir. Bu deneyi yapan öğrencinin deney yapmadan önceki hipotezi aşağıdakilerden hangisi olabilir? **(sps formulating hypothesis)**

A) Devreye bağlanan telin cinsi ampulün parlaklığını etkiler.
B) Devreye bağlanan telin uzunluğu ampulün parlaklığını etkiler.
C) Devreye bağlanan telin kalınlığı ampulün parlaklığını etkiler.
D) Devreye bağlanan ampul sayısındaki değişim ampulün parlaklığını etkiler.

**18.** Bir elektrik devresinde ampulün parlaklığı ile devreye bağlanan telin kesit alanı arasındaki ilişkiyi araştıran bir öğrencinin aşağıdaki düzeneklerden hangisini kullanması doğru olur? **(sps designing experiment)**

| | Kontrol Grubu | Deney grubu |
|---|---|---|
| A) | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet ince bakır tel | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet ince demir tel |
| B) | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet ince bakır tel | 2 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet ince bakır tel |
| C) | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet ince demir tel | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet kalın demir tel |
| D) | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet kısa demir tel | 1 adet güç kaynağı, 1 adet ampul, elektrik devresi kablosu, 1 adet uzun demir tel |

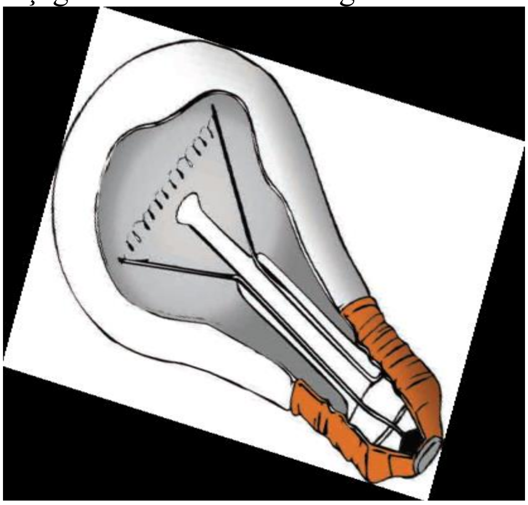
**19.** Bir öğrenci ampul parlaklığına etki eden faktörleri incelemek için kontrol grubu ve deney grubunu oluşturmak istemektedir.

| Kontrol Grubu | Deney Grubu |
|---|---|
| 1 adet güç kaynağı | 1 adet güç kaynağı |
| 1 adet ampul | 1 adet ampul |
| 1 adet 20 cm boyunda ve 1 cm eninde demir tel | 1 adet 20 cm boyunda ve 2 cm eninde demir tel |

Bu öğrenci deneyinde devreye bağlanan telin kesit alanı arttıkça ampulün parlaklığının azalacağını düşünmektedir. Deneyi yaptıktan sonra ise telin kesit alanı arttıkça ampulün parlaklığının arttığını gözlemlemiştir. Bu bilgilere göre öğrencinin hipotezinin doğruluğu hakkında aşağıdaki ifadelerden hangisini söylenebilir? **(SPS Formulting hypothesis)**

A) Hipotez yanlışlanmıştır çünkü kesit alanı arttıkça ampulün parlaklığı artmıştır.
B) Hipotez doğrulanmıştır çünkü kesit alanı arttıkça ampulün parlaklığı azalmıştır.
C) Hipotez yanlışlanmıştır çünkü deney gruplarında farklı kalınlıkta teller kullanılmıştır.
D) Hipotez doğrulanmıştır çünkü deneyler elektriği ileten maddeler ile yapılmıştır.

**20.** Bir elektrik devresi olan ampulün camı kırılıp içindeki yapı (filaman) incelendiğinde iki uç arasında ince bir tel görülür ve bu tel kıvrımlar halindedir. Aşağıdaki ifadelerden hangisi filaman ile ilgili doğru bir bilgi değildir?



A) Filamanı oluşturan telin kıvrımlı olmasının sebebi ampule uzun bir tel yerleştirmektir.
B) Filamanı oluşturan telin ince olması ampulün direncini artırır.
C) Filamanı oluşturan telin kıvrımlı olması teli daha kalın hale getirir.
D) Filamanın bağlı olduğu iki uç da elektrik enerjisini iletir.

## G. EPISTEMOLOGICAL BELIEF QUESTIONNAIRE (TURKISH FORM)

| | Kesinlikle Katılmıyorum | Katılmıyorum | Kararsızım | Katılıyorum | Kesinlikle Katılıyorum |
|---|---|---|---|---|---|
| 1. Tüm insanlar, bilim insanlarının söylediklerine inanmak zorundadır. (S) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 2. Bilimsel deneylerdeki fikirler, olayların nasıl meydana geldiğini merak edip düşünerek ortaya çıkar. (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 3. Günümüzde bazı bilimsel düşünceler, bilim insanlarının daha önce düşündüklerinden farklıdır. (D) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 4. Bir deneye başlamadan önce, deneyle ilgili bir fikrinizin olmasında yarar vardır. (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 5. Bilimsel kitaplarda yazanlara inanmak zorundasınız. (S) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 6. Bilimsel kitaplardaki bilgiler bazen değişir. (D) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 7. Bilimsel çalışmalarda düşüncelerin test edilebilmesi için birden fazla yol olabilir (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 8. Fen Bilgisi dersinde, öğretmenin söylediği herşey doğrudur (S) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 9. Bilimdeki düşünceler, konu ile ilgili kendi kendinize sorduğunuz sorulardan ve deneysel çalışmalarınızdan ortaya çıkabilir. (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 10. Bilim insanları bilim hakkında hemen hemen her şeyi bilir, yani bilinecek daha fazla bir şey kalmamıştır. (C) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 11. Bilim insanlarının bile yanıtlayamayacağı bazı sorular vardır (D) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 12. Olayların nasıl meydana geldiği hakkında yeni fikirler bulmak için deneyler yapmak, bilimsel çalışmanın önemli bir parçasıdır (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 13. Bilimsel kitaplardan okuduklarınızın doğru olduğundan emin olabilirsiniz. (S) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 14. Bilimsel bilgi her zaman doğrudur. (C) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 15. Bilimsel düşünceler bazen değişir. (D) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |

| | Kesinlikle Katılmıyorum | Katılmıyorum | Kararsızım | Katılıyorum | Kesinlikle Katılıyorum |
|---|---|---|---|---|---|
| 16. Sonuçlardan emin olmak için, deneylerin birden fazla tekrarlanmasında fayda vardır (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 17. Sadece bilim insanları , bilimde neyin doğru olduğunu kesin olarak bilirler. (S) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 18. Bilim insanının bir deneyden aldığı sonuç, o deneyin tek yanıtıdır. (C) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 19. Yeni buluşlar, bilim insanlarının doğru olarak düşündüklerini değiştirir. (D) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 20. Bilimdeki, parlak fikirler sadece bilim insanlarından değil, herhangi birinden de gelebilir (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 21. Bilim insanları bilimde neyin doğru olduğu konusunda her zaman hemfikirdirler. (C) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 22. İyi çıkarımlar, birçok farklı deneyin sonucundan elde edilen kanıtlara dayanır. (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 23. Bilim insanları, bilimde neyin doğru olduğu ile ilgili düşüncelerini bazen değiştirirler. (D) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |
| 24. Bir şeyin doğru olup olmadığını anlamak için deney yapmak iyi bir yoldur. (J) | 1❑ | 2❑ | 3❑ | 4❑ | 5❑ |

## H. TEST OF INTEGRATED PROCESS SKILLS 2 (TIPSII)

**1)** Arabaların verimliliğini inceleyen bir araştırma yapılmaktadır. Test edilen hipotez, benzine katılan katkı maddesinin arabaların verimliliğini arttırdığı yolundadır. Aynı tip beş arabaya, aynı miktarda benzin ve farklı miktarlarda katkı maddesi konulur. Arabalar benzinleri bitinceye kadar aynı yol boyunca giderler. Daha sonra her arabanın aldığı mesafe kaydedilir. Bu çalışmada arabaların verimliliği sizce nasıl ölçülür?
a. Arabaların benzinleri bitinceye kadar geçen süre ile.
b. Her arabanın gittiği mesafe ile.
c. Kullanılan benzin miktarı ile.
d. Kullanılan katkı maddesinin miktarı ile.

**2)** Bir araba üreticisi daha ekonomik arabalar yapmak istemektedir. Araştırmacılar arabanın litre başına alabileceği mesafeyi etkileyebilecek değişkenleri araştırmaktadırlar. Sizce aşağıdaki değişkenlerden hangisi arabanın litre başına alabileceği mesafeyi etkileyebilir?
a. Arabanın ağırlığı.
b. Motorun hacmi.
c. Arabanın rengi
d. a ve b.

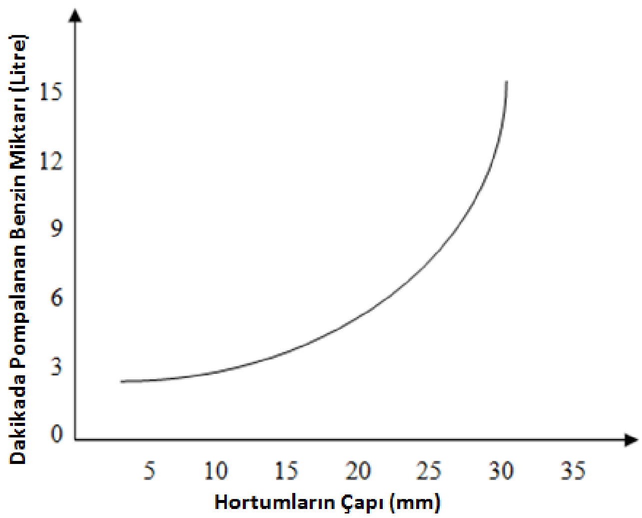
**3)** Bir polis şefi, araç kullanma hızının azaltılması ile uğraşmaktadır. Araç kullanma hızını etkileyebilecek bazı faktörler olduğunu düşünmektedir. Sürücülerin ne kadar hızlı araba kullandıklarını sizce aşağıdaki hipotezlerin hangisiyle test edilebilir?
a. Daha genç sürücülerin daha hızlı araba kullanma olasılığı yüksektir.
b. Kaza yapan arabalar ne kadar büyükse, içindeki insanların yaralanma olasılığı o kadar azdır.
c. Yollarda ne kadar çok polis ekibi olursa, kaza sayısı o kadar az olur.
d. Arabalar eskidikçe kaza yapma olasılıkları artar.

**4)** Bir fen dersinde, tekerlek genişliğinin tekerleğin daha kolay yuvarlanması üzerine etkisi araştırılmaktadır. Bir oyuncak arabaya geniş tekerlekler takılır, önce bir rampadan (eğik düzlem) aşağı bırakılır ve daha sonra düz bir zemin üzerinde gitmesi sağlanır. Deney, aynı arabaya daha dar tekerlekler takılarak tekrarlanır. Hangi tip tekerleğin daha kolay yuvarlandığı sizce nasıl ölçülür?
a. Her deneyde arabanın gittiği toplam mesafe ölçülür.
b. Rampanın (eğik düzlem) eğim açısı ölçülür.
c. Her iki deneyde kullanılan tekerlek tiplerinin genişlikleri ölçülür.
d. Her iki deneyin sonunda arabanın ağırlıkları ölçülür.

**5)** Ahmet basketbol topunun içindeki hava arttıkça, topun daha yükseğe sıçrayacağını düşünmektedir. Bu hipotezi araştırmak için, birkaç basketbol topu alır ve içlerine farklı miktarda hava pompalar. Sizce Ahmet hipotezini nasıl test etmelidir?

a. Topları aynı yükseklikten fakat değişik hızlarla yere vurur.
b. İçlerinde farklı miktarlarda hava olan topları, aynı yükseklikten yere bırakır.
c. İçlerinde aynı miktarlardaki hava olan topları, zeminle farklı açılardan yere vurur.
d. İçlerinde aynı miktarlarda hava olan topları, farklı yüksekliklerden yere bırakır.

**6)** Bir tankerden benzin almak için farklı genişlikte 5 hortum kullanılmaktadır. Her hortum için aynı pompa kullanılır. Yapılan çalışma sonunda elde edilen bulgular aşağıdaki grafikle gösterilmiştir. Size göre aşağıdakilerden hangisi değişkenler arasındaki ilişkiyi açıklamaktadır?

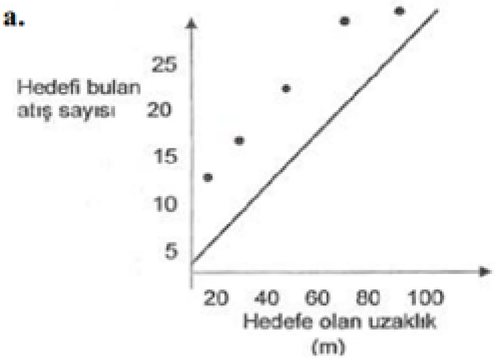



a. Hortum genişledikçe dakikada pompalanan benzin miktarı da artar.
b. Dakikada pompalanan benzin miktarı arttıkça, daha fazla zaman gerekir.
c. Hortum daraldıkça dakikada pompalanan benzin miktarı da artar.
d. Pompalanan benzin miktarı azaldıkça, hortum genişler.

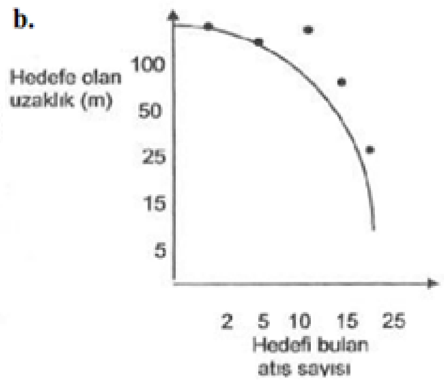
**7)** Bir hedefe çeşitli mesafelerden 25 er atış yapılır. Her mesafeden yapılan 25 atıştan hedefe isabet edenler aşağıdaki tabloda gösterilmiştir. Bu tabloya göre aşağıdaki grafiklerden hangisi çizilmelidir?

| **Mesafe (m)** | **Hedefe vuran atış sayısı** |
|---|---|
| 5 | 25 |
| 15 | 10 |
| 25 | 10 |
| 50 | 5 |
| 100 | 2 |

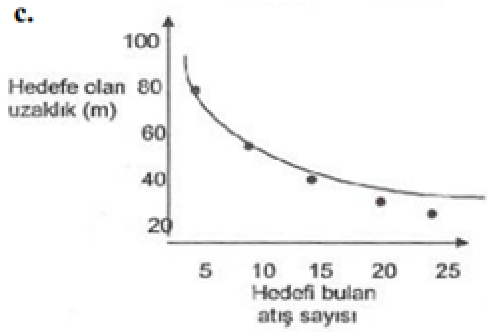
**a.**



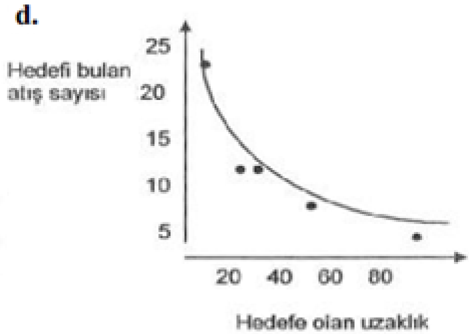

**b.**




**c.**




**d.**




***8., 9., 10. ve 11. soruları aşağıdaki bilgiye göre cevaplayınız.*** *Ayşe, güneşin karaları ve denizleri aynı derecede ısıtıp ısıtmadığını merak etmektedir. Bir araştırma yapmaya karar verir ve aynı büyüklükte iki kova alır. Bunlardan birini toprakla, diğerini de su ile doldurur ve aynı miktarda güneş ışığı alacak şekilde bir yere koyar. Günün 8.00-18.00 saatleri arasında, her saat başı sıcaklıklarını ölçer.*

**8)** Sizce araştırmada aşağıdaki hipotezlerden hangisi test edilmiştir?
a. Toprak ve su ne kadar çok güneş ışığı alırlarsa, o kadar ısınırlar.
b. Toprak ve su güneş altında ne kadar fazla kalırlarsa, o kadar çok ısınırlar.
c. Güneş farklı maddeleri farklı derecelerde ısıtır.
d. Günün farklı saatlerinde güneşin yaydığı ısı da farklı olur.

**9)** Sizce araştırmada aşağıdaki değişkenlerden hangisi sabit tutulmuştur?
a. Kovadaki suyun cinsi.
b. Toprak ve suyun sıcaklığı.
c. Kovalara koyulan maddenin türü.
d. Her bir kovanın güneş altında kalma süresi.

**10)** Sizce araştırmada ölçülen değişken hangisidir?
a. Kovadaki suyun cinsi.
b. Toprak ve suyun sıcaklığı.
c. Kovalara koyulan maddenin türü.
d. Her bir kovanın güneş altında kalma süresi.

**11)** Sizce araştırmada değiştirilen değişken hangisidir?
a. Kovadaki suyun cinsi.
b. Toprak ve suyun sıcaklığı.
c. Kovalara koyulan maddenin türü.
d. Her bir kovanın güneş altında kalma süresi.

***12., 13., 14. ve 15. soruları aşağıdaki bilgiye göre cevaplayınız. Murat****, suyun sıcaklığının, su içinde çözünebilecek şeker miktarını etkileyip etkilemediğini araştırmak ister. Birbirinin aynı dört bardağın her birine 50 mililitre su koyar. Bardaklardan birisine $0^0$ C de, diğerlerine de sırayla $50^0$ C, $75^0$ C ve $95^0$ C sıcaklıkta su koyar. Daha sonra her bir bardağa çözünebileceği kadar şeker koyar ve karıştırır.*

**12)** Bu araştırmada sizce test edilen hipotez hangisi olabilir?
a. Şeker ne kadar çok suya karıştırılırsa o kadar çok çözünür.
b. Ne kadar çok şeker çözünürse, su o kadar tatlı olur.
c. Sıcaklık ne kadar yüksek olursa, çözünen şekerin miktarı da o kadar fazla olur.
d. Kullanılan suyun miktarı arttıkça sıcaklığı da artar.

**13)** Bu araştırmada sizce sabit tutulan değişken hangisidir?
a. Her bardaktaki çözünen şeker miktarı.
b. Her bardağa konulan su miktarı.
c. Bardakların sayısı.
d. Suyun sıcaklığı.

**14)** Sizce araştırmanın ölçülen değişkeni hangisidir?
a. Her bardaktaki çözünen şeker miktarı.
b. Her bardağa konulan su miktarı.
c. Bardakların sayısı.
d. Suyun sıcaklığı.

**15)** Sizce araştırmadaki değiştirilen değişken hangisidir?
a. Her bardaktaki çözünen şeker miktarı.
b. Her bardağa konulan su miktarı.
c. Bardakların sayısı.
d. Suyun sıcaklığı.

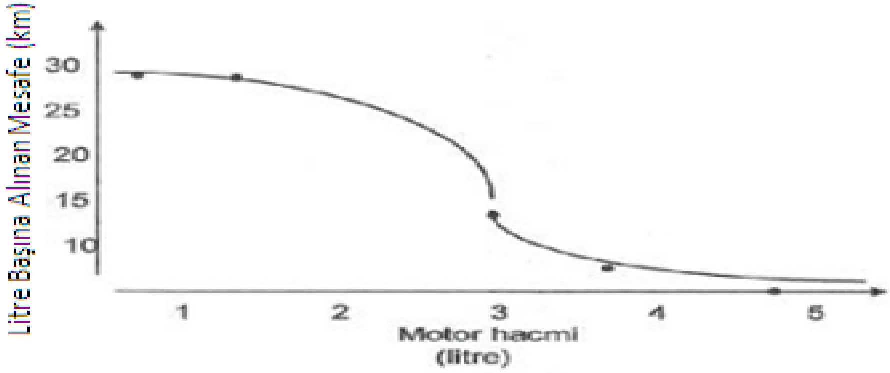
**16)** Bir bahçıvan domateslerinin çabuk filizlenmesini istemektedir. Değişik birkaç alana domates tohumu eker. Hipotezi, tohumlar ne kadar çok sulanırsa, o kadar çabuk filizleneceğidir. Sizce bu hipotezi nasıl test eder?
a. Farklı miktarlarda sulanan tohumların kaç günde filizleneceğine bakar.
b. Her sulamadan bir gün sonra domates bitkisinin boyunu ölçer.
c. Farklı alanlardaki bitkilere verilen su miktarını ölçer.
d. Her alana ektiği tohum sayısına bakar.

**17)** Ahmet, buz parçacıklarının erime süresini etkileyen faktörleri merak etmektedir. Buz parçalarının büyüklüğü, odanın sıcaklığı ve buz parçalarının şekli gibi faktörlerin erime süresini etkileyebileceğini düşünür. Daha sonra şu hipotezi sınamaya karar verir. Buz parçalarının şekli erime süresini etkiler. Sizce Ahmet bu hipotezi sınamak için aşağıdaki deney tasarımlarının hangisini uygulamalıdır?
a. Her biri farklı şekil ve ağırlıkta beş buz parçası alınır. Bunlar aynı sıcaklıkta, benzer beş kabın içine ayrı ayrı konur ve erime süreleri izlenir.
b. Her biri aynı şekilde fakat farklı ağırlıkta beş buz parçası alınır. Bunlar aynı sıcaklıkta benzer beş kabın içine ayrı ayrı konur ve erime süreleri izlenir.
c. Her biri aynı ağırlıkta fakat farklı şekillerde beş buz parçası alınır. Bunlar aynı sıcaklıkta benzer beş kabın içine ayrı ayrı konur ve erime süreleri izlenir.
d. Her biri aynı ağırlıkta fakat farklı şekillerde beş buz parçası alınır. Bunlar farklı sıcaklıkta benzer beş kabın içine ayrı ayrı konur ve erime süreleri izlenir.

**18)** Bir biyolog şu hipotezi test etmek ister; Farelere ne kadar çok vitamin verilirse o kadar hızlı büyürler. Biyolog farelerin büyüme hızını sizce nasıl ölçebilir?
a. Farelerin hızını ölçer.
b. Farelerin, günlük uyumadan durabildikleri süreyi ölçer.
c. Her gün fareleri tartar.
d. Her gün farelerin yiyeceği vitaminleri tartar.

**19)** Öğrenciler, şekerin suda çözünme süresini etkileyebilecek değişkenleri düşünmektedirler. Suyun sıcaklığını, şekerin ve suyun miktarlarını değişken olarak saptarlar. Öğrenciler, şekerin suda çözünme süresini sizce aşağıdaki hipotezlerden hangisiyle sınayabilir?
a. Daha fazla şekeri çözmek için daha fazla su gereklidir.
b. Su soğudukça, şekeri çözebilmek için daha fazla karıştırmak gerekir.
c. Su ne kadar sıcaksa, o kadar çok şeker çözünecektir.
d. Su ısındıkça şeker daha uzun sürede çözünür.

**20)** Bir araştırma grubu, değişik hacimli motorları olan arabaların verimliliğini ölçer. Elde edilen sonuçların grafiği aşağıdaki gibidir:




Sizce aşağıdakilerden hangisi değişkenler arasındaki ilişkiyi gösterir?
a. Motor ne kadar büyükse, bir litre benzinle gidilen mesafe de o kadar uzun olur.
b. Bir litre benzinle gidilen mesafe ne kadar az olursa, arabanın motoru o kadar küçük demektir.
c. Motor küçüldükçe, arabanın bir litre benzinle gidebileceği mesafe artar.
d. Bir litre benzinle gidilen mesafe ne kadar uzun olursa, arabanın motoru o kadar büyük demektir.

***21., 22., 23. ve 24. soruları aşağıdaki bilgiye göre cevaplayınız***. *Toprağa karıştırılan yaprakların domates üretimine etkisi araştırılmaktadır. Araştırmada dört büyük saksıya aynı miktarda ve tipte toprak konulmuştur. Fakat birinci saksıdaki toprağa 15 kg., ikinciye 10 kg., üçüncüye ise 5 kg. çürümüş yaprak karıştırılmıştır. Dördüncü saksıdaki toprağa ise hiç çürümüş yaprak karıştırılmamıştır. Daha sonra bu saksılara domates ekilmiştir. Bütün saksılar güneşe konmuş ve aynı miktarda sulanmıştır. Her saksıdan elde edilen domates tartılmış ve kaydedilmiştir.*

**21)** Bu araştırmada sizce test edilen hipotez hangisidir?
a. Bitkiler güneşten ne kadar çok ışık alırlarsa, o kadar fazla domates verirler.
b. Saksılar ne kadar büyük olursa, karıştırılan yaprak miktarı o kadar fazla olur.
c. Saksılar ne kadar çok sulanırsa, içlerindeki yapraklar o kadar çabuk çürür.
d. Toprağa ne kadar çok çürük yaprak karıştırılırsa, o kadar fazla domates elde edilir.

**22)** Sizce bu araştırmada sabit tutulan değişken hangisidir?
a. Her saksıdan elde edilen domates miktarı.
b. Saksılara karıştırılan yaprak miktarı.
c. Saksılardaki toprak miktarı.
d. Çürümüş yaprak karıştırılan saksı sayısı.

**23)** Sizce araştırmada ölçülen değişken hangisidir?
a. Her saksıdan elde edilen domates miktarı.
b. Saksılara karıştırılan yaprak miktarı.
c. Saksılardaki toprak miktarı.
d. Çürümüş yaprak karıştırılan saksı sayısı.

**24)** Sizce araştırmada değiştirilen değişken hangisidir?
a. Her saksıdan elde edilen domates miktarı.
b. Saksılara karıştırılan yaprak miktarı.
c. Saksılardaki toprak miktarı.
d. Çürümüş yaprak karıştırılan saksı sayısı.

**25)** Sibel, akvaryumdaki balıkların bazen çok hareketli bazen ise durgun olduklarını gözler. Balıkların hareketliliğini etkileyen faktörleri merak eder. Sizce balıkların hareketliliğini etkileyen faktörleri hangi hipotezle sınayabilir?
a. Balıklara ne kadar çok yem verilirse, o kadar çok yeme ihtiyaçları vardır.
b. Balıklar ne kadar hareketli olursa o kadar çok yeme ihtiyaçları vardır.
c. Su da ne kadar çok oksijen varsa, balıklar o kadar iri olur.
d. Akvaryum ne kadar çok ışık alırsa, balıklar o kadar hareketli olur.

**26)** Murat Bey'in evinde birçok elektrikli alet vardır. Fazla gelen elektrik faturaları dikkatini çeker. Kullanılan elektrik miktarını etkileyen faktörleri araştırmaya karar verir. Sizce aşağıdaki değişkenlerden hangisi kullanılan elektrik enerjisi miktarını etkileyebilir?
a. TV nin açık kaldığı süre.
b. Elektrik sayacının yeri.
c. Çamaşır makinesinin kullanma sıklığı.
d. a. ve c.

Test bitmiştir teşekkür ederiz ☺

## I. THIRD WEEK EVIDENCE CARDS

| Yalıtım Malzemeleri | Yanma Özelliği | Kullanıldığı Yerler | Kullanım Ömrü |
|---|---|---|---|
| Plastik köpük | Alev alır. | Dış ve iç duvar | Uzun ömürlüdür. |
| Ahşap | Alev alır. | Dış ve iç döşeme | Kısa ömürlüdür. |
| Taş yünü | Yanmaz. | Tavan, iç ve dış duvar | Uzun ömürlüdür. |
| Katran | Alev alır. | Tavan | Kısa ömürlüdür. |
| Cam yünü | Zor alev alır. | Tavan, iç ve dış duvar, güneş paneli, tesisat boruları | Uzun ömürlüdür. |
| Silikon yünü | Zor alev alır. | Dış cephe | Uzun ömürlüdür. |

| Kullanılan madde | Birim Fiyatı |
|---|---|
| Plastik Köpük | 35 TL |
| Ahşap | 180 TL |
| Taş Yünü | 60 TL |
| Katran | 20 TL |
| Cam Yünü | 40 TL |
| Silikon Yünü | 20 TL |

### Isı Yalıtım Malzemeleri

Isı katılarda en hızlı, gazlarda ise en yavaş yayılır. Bina yalıtımındaki en önemli nokta, ısının yavaş yayılmasını sağlamaktır. Buna göre bina yalıtımında hava boşlukları bulunan malzemeler kullanmak önemlidir. Havayı oluşturan tanecikler arasındaki boşluk miktarı çok fazladır. Bu nedenle havayı oluşturan tanecikler arasındaki ısı iletimi çok yavaş olur.

İçinde tanecik bulunmayan yalıtım malzemesi, ısı akışının gerçekleşmesini engellemiş olur. Bu şekilde içindeki havası boşaltılmış ortamlara **vakum** denir. Binalarda yalıtım yapılmasıyla %25-50 oranında daha az yakıt kullanılarak aynı ısınma sağlanabilir.

**1. Plastik Köpük:**

Plastik köpüğün (strafor köpük) yapısında da hava boşlukları bulunur. Bu nedenle plastik köpük ısı yalıtkanı olarak kullanılır ve ısı yalıtımı sağlanır.

**2. Katran**

Katran günümüzde sıklıkla kullanılan ağaç, kömür gibi maddelerin öğütülmesiyle oluşan yağa denir.  Renk olarak koyu renkte ve keskin kokuludur.

**3. Cam Yünü**

Sıcağa ve rutubete maruz kalması halinde dahi, boyutlarında bir değişme olmaz. Zamanla bozulmaz, çürümez, böcekler tarafından tahrip edilmez. Eritilmiş camdan elde edilir, bükülebilir, ateşe dayanıklıdır.

**4. Silikon yünü**

İçinde silisyum (Si) ve oksijen (O) atomları bulunan maddelerin ortak adıdır. Çeşitli malzemelerin su geçirgenliğini azaltmakta kullanılır. Geniş bir sıcaklık aralığında esnekliği korur.

**5. Taş Yünü**

Taş yünü, volkanik taşlardan tedarik edilen minerallerin çok yüksek sıcaklıklarda eritilmesi ile üretilir. Yangına karşı çok dayanıklı ve su iticiliğiyle iyi bir yalıtım malzemesidir.

**6. Ahşap Yünü**

Uzun lifler haline getirilmiş ahşap talaşının yüksek sıcaklıkta baskılanmasıyla üretilen yalıtım ürünüdür. Çok az kullanım alanı bulan ahşap yünü doğal yapısı gereği çevre dostudur. Güneşin ışınlarından çok etkilenmese de organik bir madde olduğu için böcek ve organizmalardan zarar görebilir.

## J. REFORMED TEACHING OBSERVATION PROTOCOL (RTOP)

| No | Item | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|
| **1.** | The instructional strategies and activities respected students' prior knowledge and the preconceptions inherent therein. | | | | | |
| **2.** | The focus and the direction of the lesson were often determined by ideas originating with students. | | | | | |
| **3.** | Students were actively engaged in thought-provoking activity that often involved the critical assessment of procedures. | | | | | |
| **4.** | Students were reflective about their learning. | | | | | |
| **5.** | Intellectual rigor, constructive criticism, and the challenging of ideas were valued. | | | | | |
| **6.** | Students were involved in the communication of their ideas to others using a variety of means and media. | | | | | |
| **7.** | The teachers' questions triggered divergent modes of thinking. | | | | | |
| **8.** | There was a high proportion of student talk and a significant amount of it occurred between and among students. | | | | | |
| **9.** | Student questions and comments often determined the focus and direction of classroom discourse. | | | | | |
| **10.** | Active participation of students was encouraged and valued. | | | | | |
| **11.** | Students were encouraged to generate conjectures, alternative solution strategies, and/or different ways of interpreting evidence. | | | | | |
| **12.** | The teacher acted as a resource person, working to support and enhance student investigations. | | | | | |
| **13.** | The metaphor "teacher as listener" was very characteristic of this classroom. | | | | | |

## K. METU ETHICAL COMMITTEE APPROVAL

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
Sayı: 28620816 / 175
www.ueam.metu.edu.tr

05 NİSAN 2018

Konu: Değerlendirme Sonucu

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi: İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Prof.Dr. Semra SUNGUR

Danışmanlığını yaptığınız doktora öğrencisi Mehmet ŞEN'in "***Argümantasyon Tabanlı Bilim Öğrenme Yöntemi Aracılığıyla Öğrencilerin Fen Okuryazarlığının Geliştirilmesi***" başlıklı araştırması İnsan Araştırmaları Etik Kurulu tarafından uygun görülerek gerekli onay **2018-FEN-014** protokol numarası ile **06.04.2018 - 30.09.2019** tarihleri arasında geçerli olmak üzere verilmiştir.

Bilgilerinize saygılarımla sunarım.

Prof. Dr. Ş. Halil TURAN
Başkan V

Prof. Dr. Ayhan SOL
Üye

Prof. Dr. Ayhan Gürbüz DEMİR
Üye

Doç. Dr. Yaşar KONDAKÇI
Üye

Doç. Dr. Zana ÇITAK
Üye

Doç. Dr. Emre SELÇUK
Üye

Dr. Öğr. Üyesi Pınar KAYGAN
Üye

## L. MINISTRY OF NATIONAL EDUCATION ETHICAL COMMITTEE APPROVAL

T.C.
ANKARA VALİLİĞİ
Milli Eğitim Müdürlüğü

Sayı : 14588481-605.99-E.13832137 26.07.2018
Konu : Araştırma İzni

ORTA DOĞU TEKNİK ÜNİVERSİTESİNE
(Öğrenci İşleri Daire Başkanlığı)

İlgi: a) MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 2017/25 nolu Genelgesi.
b) 27/06/2018 Tarihli ve E.17 sayılı yazınız.

Üniversiteniz İlköğretim Anabilim Dalı Doktora Programı öğrencisi Mehmet ŞEN'in **"Argümantasyon Tabanlı Bilim Öğrenme Yöntemi Aracılığıyla Öğrencilerin Fen Okuryazarlığın Gelişirilmesi"** konulu uygulama talebi Müdürlüğümüzce uygun görülmüş ve uygulamanın yapılacağı İlçe Milli Eğitim Müdürlüğüne bilgi verilmiştir.

Görüşme formunun (37 sayfa) araştırmacı tarafından uygulama yapılacak sayıda çoğaltılması ve çalışmanın bitiminde bir örneğinin (cd ortamında) Müdürlüğümüz Strateji Geliştirme (1) Şubesine gönderilmesini rica ederim.

Vefa BARDAKCI
Vali a.
Milli Eğitim Müdürü

Güvenli Elektronik İmzalı
Aslı İle Aynıdır.
23../2../2018

Adres: Emniyet Mah. Alparslan Türkeş Cad. 4/A
Yenimahalle/ANKARA
Elektronik Ağ: www.meb.gov.tr
e-posta: istatistik06@meb.gov.tr

Bilgi için: D. KARAGÜZEL
Tel: 0 (312) 221 02 17
Faks: 0 (312) 221 02 16

Bu evrak güvenli elektronik imza ile imzalanmıştır. https://evraksorgu.meb.gov.tr adresinden a2be-b781-374e-92e8-61bb kodu ile teyit edilebilir.

## M. PARTICIPATION FORM

Bu çalışma Argümantasyon Tabanlı Bilim Öğrenme (ATBÖ) yönteminin ortaokul öğrencilerinin fen okuyazarlık düzeyine olan etkisini incelemektedir. Çalışmaya katılım gönüllülük esasına dayalıdır. Bu çalışma kapsamında herhangi bir kimlik bilgisi istenmemektedir. Maddelere verdiğiniz cevaplar ve video kayıtları tamamen gizli tutulacak ve veriler araştırmacı tarafından değerlendirilecektir. Elde edilen veriler bilimsel amaçla kullanılacaktır.

ATBÖ uygulamasının katılımcılara herhangi bir zararı bulunmamaktadır. Çalışmayı yarıda bırakabilir veya araştırmacıya konu ile ilgili sorular sorabilir, fikirlerinizi araştırmacı ile paylaşabilirsiniz.

Çalışmaya katkı sağladığınız için teşekkür ederiz. Çalışma hakkında daha fazla bilgi sahibi olmak açısından Orta Doğu Teknik Üniversitesi Matematik ve Fen Eğitimi Ana Bilim Dalı Araştırma Görevlisi Mehmet Şen ile msen@metu.edu.tr adresi üzerinden iletişime geçebilirsiniz.

## N. PARENT PREMISSION FORM

Sevgili Anne-Babalar;

Ben Orta Doğu Teknik Üniversitesi Matematik ve Fen Bilimleri Eğitimi Bölümü'nden araştırma görevlisi Mehmet Şen. Doktora tez çalışmam kapsamında "Argümantasyon Tabanlı Bilim Öğrenme Yöntemi Aracılığıyla Öğrencilerin Fen Okuryazarlığının Geliştirilmesi" isimli çalışmayı yapmayı planlamaktayım.

Araştırmanın amacı Argümantasyon Tabanlı Bilim Öğrenme (ATBÖ) Yaklaşımının öğrencilerin fen okuryazarlık bileşenleri üzerine etkisini ölçmektir. ATBÖ yönteminde öğrenciler bireysel, grup halinde ve tüm sınıf olarak araştırma soruları oluşturacak, deney malzemeleri ile araştırma sorularını test edecek, sonuçları yorumlayacak ve sonuçlar üzerinden bilimsel argümanlar ve tartışmalar ortaya koyacaklardır. Ayrıca, sözlü argümanlarını yazılı raporlarla destekleyeceklerdir.

Çalışmanın amacını gerçekleştirebilmek için çocuklarınızın bazı anketleri doldurmasına ve ders uygulamasının video ile kayıt altına alınmasına ihtiyaç duymaktayız. Katılmasına izin verdiğiniz takdirde çocuğunuz anketi okulda ders saatinde dolduracaktır. Sizden çocuğunuzun katılımcı olmasıyla ilgili izin istediğimiz gibi, çalışmaya başlamadan çocuğunuzdan da sözlü olarak katılımıyla ilgili rızası mutlaka alınacaktır.

Çocuğunuzun dolduracağı anketlerde cevaplar kesinlikle gizli tutulacak, video kayıtları hiçkimse ile paylaşılmayacak ve çalışmadan elde edilen sonuçlar sadece bilimsel araştırma amacıyla kullanılacaktır. Çocuğunuzun çalışmaya katılması sonucunda fen bilgisi dersinin genel kazanımı olan fen okuryazarlığına ilişkin çocuğunuzun çalışma öncesindeki ve sonucundaki içerik alan bilgisini (konu alan bilgisi), bilimsel uygulama kapasitesini, argümantasyon becerilerini, epistemolojik inançlarını (bilgiye yönelik inançlar) ve bilimsel süreç becerilerini görebilirsiniz.

Çocuğunuzun çalışmaya katılmasının onun psikolojik gelişimine olumsuz etkisi olmayacağından emin olabilirsiniz. Yine de, bu formu imzaladıktan sonra çocuğunuz katılımcılıktan ayrılma hakkına sahiptir. Herhangi bir uygulama ile ilgili bir nedenden ötürü çocuğunuz kendisini rahatsız hissettiğini belirtirse, ya da kendi belirtmese de araştırmacı çocuğun rahatsız olduğunu öngörürse, çalışmaya derhal son verilecektir. Araştırmayla ilgili sorularınızı aşağıdaki e-posta adresini kullanarak bana yöneltebilirsiniz.

Saygılarımla,

Arş Gör. Mehmet Şen
Matematik ve Fen Bilimleri Eğitimi Bölümü
Orta Doğu Teknik Üniversitesi, ANKARA
e-posta: msen@metu.edu.tr

Lütfen bu araştırmaya çocuğunuzun katılması konusundaki tercihinizi aşağıdaki seçeneklerden size en uygun gelenin altına imzanızı atarak belirtiniz ve bu formu çocuğunuzla okula geri gönderiniz.

A) Bu araştırmaya çocuğum .....................................’nın da katılımcı olmasına izin veriyorum. Çalışmayı istediği zaman yarıda kesip bırakabileceğini biliyorum ve verdiği bilgilerin bilimsel amaçlı olarak kullanılmasını kabul ediyorum.

Baba Adı Soyadı......................................... Anne Adı Soyadı....................................

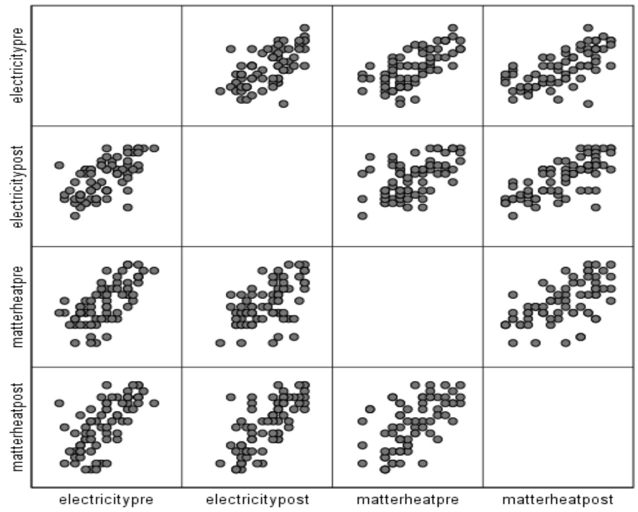
İmza.......................................................... İmza.......................................................

B) Bu çalışmada çocuğum...................................’nın katılımcı olmasına izin vermiyorum.

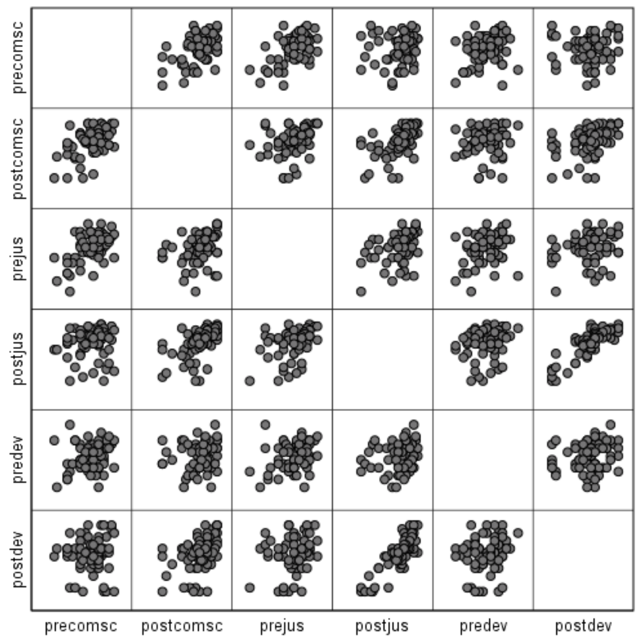
Baba Adı Soyadı......................................... Anne Adı Soyadı....................................

İmza.......................................................... İmza.......................................................

**O. MATRIX SCATTERPLOTS FOR LINEARITY ASSUMPTION**




Matrix Scatterplot for Linearity Assumption regarding content knowledge changes




Matrix Scatterplot for Linearity Assumption regarding epistemological beliefs changes

## P. CURRICULUM VITAE

**<u>Areas of Research Interest</u>**

Argumentation in Science Education
Pedogogical Content Knowledge (PCK)

**<u>Educational Background</u>**

| | | |
|---|---|---|
| PhD. | 2021 | Elementary Education<br>Middle East Technical University, Ankara |
| M.S. | 2014 | Math and Science Education<br>Middle East Technical University, Ankara |
| B. S. | 2011 | Elementary Science Education<br>Middle East Technical University, Ankara |

**<u>Appointments</u>**

2012- 2021 Research and Teaching Assistant, Middle East Technical University, Ankara

**<u>Master Thesis:</u>**

**Sen, M.** (2014). *A study on science teachers' pedagogical content knowledge and content knowledge regarding cell division* (Unpublished master's thesis). Middle East Technical University, Ankara.

**<u>Peer-Reviewed Journal Papers</u>**

**Sen, M.,** Öztekin, C., & Demirdöğen, B. (2018). Impact of content knowledge on pedagogical content knowledge in the context of cell division. *Journal of Science Teacher Education*, 29(2), 102-127.

**Sen, M.,** & Öztekin, C. (2019). Interaction among contextual knowledge and pedagogical content knowledge: Sociocultural perspective. *Education and Science*, 44(198), 57-97.

**Peer-Reviewed Conference Papers**

Yeşilyurt, E. & **Sen, M.** (2013, April). *Perceptions of university students' paranormal beliefs.* Poster session presented at the meeting of International Conference On Innovation and Challenges in Education 2013 (CICE ,2013), Kütahya, Turkey.

Kıran, D., & **Sen, M.** (2014). In-service science teacher profiles from the eyes of pre-service science teachers: What did they observe? *Proceedings of International Conference on Education in Mathematics, Science & Technology (ICEMST 2014)*, 681-683.

**Sen, M.** & Oztekin, C. (2014, September). *Hücre bölünmesi konusunda sürece yönelik konu alan bilgisi ve pedagojik alan bilgisinin incelenmesi*. Paper session presented at the meeting on 11.Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi (XI. UFBMEK), Adana.

**Sen, M.** (2015, April). *A beginning science teacher's pck and content knowledge in cell division topics.* Paper session presented at the meeting on The International Organization for Science and Technology Education (IOSTE), Istanbul.

Aydın, G. C., Evren E., Atakan, I., **Sen, M.,** Yılmaz B., Pirgon E.,... Ebren, E. (2015, June). *Delphi technique as a graduate course activity: Elementary science teachers' TPACK competencies*. Paper session presented at the meeting on ERPA International Congresses on Education, Athens (Greece).

**Sen, M.** & Oztekin, C. (2015, September). *A study on experienced science teachers' pedagogical content knowledge regarding cell division.* Paper session presented at the meeting on The European Conference on Educational Research (ECER), Budapest, Hungary.

**Sen, M.** (2015, September). *İlköğretim 3-8. Sınıflarda kullanılan fen bilimleri ders ve çalışma kitaplarında yer alan fizik konularını içeren çoktan seçmeli soruların geçerlilik yönünden incelenmesi.* Paper session presented at the meeting on II. Ulusal Fizik Eğitimi Kongresi (UFEK), Ankara, Turkey.

Demirci, S. & **Sen, M.** (2015, September). *Fen bilgisi öğretmen adaylarının yoğunluk konusunda kavram yanılgılarının belirlenmesi.* Paper session presented at the meeting on II. Ulusal Fizik Eğitimi Kongresi (UFEK), Ankara, Turkey.

**Sen, M.** & Öztekin, C. **(**2016, June**)**. *Teaching density: To what extent science teachers know about instructional strategies and student understanding?* Paper presented IIIrd International Eurasian Educational Research Congress/ 3. Uluslararası Avrasya Eğitim Araştırmaları *Kongresi.* Muğla Sıtkı Koçman Üniversitesi, Muğla/Türkiye 31 Mayıs-03 Haziran

Demirci, S. & **Sen, M.** (2016, April). *Pre-service science teachers' misconceptions about density and buoyancy.* Paper presented at the 89th National Association for Research in Science Teaching (NARST) Annual International Conference. Baltimore, MD, USA 14-17 April

**Sen, M.** & Öztekin, C. (2016, September). *Fen bilimleri öğretmenleri neyi öğretiyor, neyi ölçüyor?* Paper session presented at the meeting on 12.Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi (12. UFBMEK), Trabzon.

Kaplan, G., **Sen, M.** & Kıran, D. (2016, September). *Ortaokul öğrencilerinin matematik ve fen bilimleri derslerindeki başarı hedef yönelimlerinin sınıf seviyelerine göre karşılaştırılması*. Paper session presented at the meeting on 12.Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi (12. UFBMEK), Trabzon.

Özdemir, M., **Sen, M.**, & Öztekin, C. (2017, April). *Students' understandings and feelings about sustainable development.* Paper presented at the 26th International Conference on Educational Sciences/26. Uluslararası Eğitim Bilimleri Kongresi, Antalya/Türkiye 20-23 Nisan

Demirci, S., Kaplan, G., Cengizoğlu, S., & **Sen, M.** (2017, April). *Pre-service teachers' understandings of sustainability based on ontological entities.* Paper presented at the 26th International Conference on Educational Sciences/26. Uluslararası Eğitim Bilimleri Kongresi, Antalya/Türkiye 20-23 Nisan

Aydın, R., **Sen, M.** & Öztekin, C (2017, May). *To what extent are science teachers ready to teach sustainable development? A case study for two experienced teachers pedagogical content concerns.* Paper presented at the ERPA International Congresses on Education 2017, Budapest/Hungary, 18-21 May 2017. (18-21 Mayıs 2017 ERPA uluslararası Eğitim Kongresi ERPA).

Tiras, I., Öztekin, C. & **Sen, M.** (August, 2017). *Preservice elementary science teachers' pedagogical content knowledge for sustainable development.* Paper presented at the European Science Education Research Association (ESERA) 2017, Dublin/Ireland, 21-25 August 2017.

**Sen, M.**, Sungur, S. & Öztekin, C. (June, 2018). *Improving middle school students' scientific literacy through science writing heuristic (SWH) approach.* Paper presented at the ERPA International Congresses on Education 2018, Istanbul/Turkey, 28 June-1 July 2018. (28 Haziran-1 Temmuz 2018 ERPA uluslararası Eğitim Kongresi ERPA).

**Sen, M.**, Sungur, S. & Öztekin, C. (June, 2018). *Adaptation of global scientific literacy questionnaire in turkish context: Findings of pilot study.* Paper presented at the ERPA International Congresses on Education 2018,

Istanbul/Turkey, 28 June-1 July 2018. (28 Haziran-1 Temmuz 2018 ERPA uluslararası Eğitim Kongresi ERPA).

**Sen, M.**, & Öztekin, C. (September, 2018). *Examination of middle school science teachers' contextual knowledge and pedagogical content knowledge in density topic.* Paper presented at the ECER European Educational Research Association 2018, Bolzano/ Italy, 3-7 September 2018.

**Sen, M.**, Güler, F., Özdemir, M., & Öztekin, C. (September, 2018). *Uncovering teacher candidates' sustainability consciousness level.* Paper presented at the ECER European Educational Research Association 2018, Bolzano/ Italy, 3-7 September 2018.

Güler, F., **Sen, M.**, Özdemir, M., & Öztekin, C. (September, 2018). *Adaptation and validation of the sustainability consciousness questionnaire for teacher candidates.* Paper presented at the ECER European Educational Research Association 2018, Bolzano/ Italy, 3-7 September 2018.

**Scientific Research Projects**

A Study on Pre-service Teachers Evolution Literacy Level, METU-BAP-05-06-2013-002 (Researcher).

A Study on Science Teachers Contextual Knowledge and Pedagogical Content Knowledge in Density Topic. METU-BAP-05-06-2016-001 (Researcher).

Adaptation of Global Scientific Literacy Scale in Turkish: Validity and Reliability Study. METU-BAP--05-07-2017-001 (Researcher)

Investigation of Middle School Students and Pre-service Teachers' Sustainability Consciousness. METU- BAP- 05-06-2017-001 (Researcher)

Engineering applications for sustainable development: Green science, technology, mathematics and engineering education (Green STEM). The Scientific and

Technological Research Council of Turkey (TÜBİTAK) 4004 PROJECT, PROJECT CODE: 218B020 METU, ANKARA, 16-22 June 2019 (Guide)

**Courses Taught (Teaching Assistant)**

**Graduate Courses:**

ESME 543 Skills on Environmental Issue Investigation

ESME 560 Analysis of Research in Elementary Math. & Science Educ.

ELE 603 Advanced Educational Research

MSE 502 Research Methods and Ethics

**Undergraduate Courses:**

ELE 343 Methods of Teaching Sciences I

ELE 344 Methods of Teaching Science II

ELE 331 Laboratory Applications in Science I

ELE 332 Laboratory Applications in Science II

ELE 419 School Experience

ELE 420 Practice Teaching in Elementary Education

ELE 310 Community Service

ELE 240 Probability and Statistics

ELE 225 Measurement and Assessment

ECE 120 Anatomy and Physiology

ELE 440 Science, Technology and Society

ECE 250 Basic Science

ELE 475 Climate Change Education for Sustainability

ELE 329 Instructional Technology and Material Development

ELE 221 Instructional Principles and Methods

# Q. TURKISH SUMMARY/ TÜRKÇE ÖZET

## GİRİŞ

Fen öğretiminin temel amacı bilimsel okuryazarlığı gelişmiş bireyler yetiştirmektir (Roberts, 2007). Fen okuryazarlığına ilişkin farklı yaklaşımlar olmasına karşın, vizyon-1 yaklaşımına göre bilim insanlarının sahip olduğu özellikler bilimsel okuryazarlık için ölçüt olarak kabul edilmiştir. Buna göre; bilimsel okuryazarlığı gelişmiş öğrenciler geleceğin olası bilim insanlarıdır. Vizyon-1 yaklaşımı bilimin hem süreç hemde ürün (ör; alan bilgisi) özellikleri ile ilgilidir (Roberts, 2007). Bu çalışmada vizyon-1 yaklaşımı benimsenmiştir. Vizyon-1 yaklaşımı benimsendiği için, bilimle doğrudan ilişkili olan değişkenler seçilmiş ve bu değişkenlerin bilimsel okuryazarlığı temsil ettiği düşünülmüştür. Bilimle doğrudan ilişkili olduğu düşünülen ve bilimsel okuryazarlığı temsil ettiği düşünülen değişkenler bu çalışmada 6. sınıf öğrencilerinin içerik alan bilgisi, epistemolojik inançları ve bilimsel süreç becerileridir. İçerik alan bilgisi bilimsel çabanın ürünüdür ve bu çalışmada içerik alan bilgisi fizik, kimya, biyoloji gibi disiplinlerin akademik içeriğini temsil etmektedir (Carlson & Daehler, 2019).

Bilimsel okuryazarlık aynı zamanda bilimin işleyiş sürecini de içerir, buna göre bilimsel okuryazar bir birey bilimsel bir dilde konuşur, yazar ve bilimsel etkinliklere katılır. Bilimsel etkinliklere katıldığı için de bilimsel bilginin nasıl üretildiği konusunda ve bilimsel bilginin özellikleri konusunda bilgi sahibi olur (Jimenez-Aleixandre, & Erduran, 2008). Bu durum kişinin epistemolojik inançları ile ilişkilidir ve bu nedenle epistemolojik inançlar bilimsel okuryazarlığın bir göstergesi olarak bu çalışmada kabul edilmiştir. Bu çalışmada epistemolojik inançlar bilimsel bilginin kaynağı, kesinliği, gelişimi ve gerekçelendirmesini içeren inanç sistemi olarak tanımlanmıştır (Conley, Pintrich, Vekiri, & Harrison, 2004). Bu çalışmada yer alan bilimin işleyiş süreci ile ilgili olduğu düşünülen bir başka bilimsel okuryazarlık boyutu ise bilimsel süreç becerileridir. Çünkü bilimsel bilgi üretilirken kişilerin bilimsel süreç becerilerini kullanması gerekmektedir (Harlen, 1999). Bilimsel süreç becerileri bilim insanlarının düşünme sırasında kullandıkları özelliklerin sınıflandırılmasıyla ortaya konan bilim ile ilişkili becerilerdir (Padilla, 1990; Sanderson & Kratochvil, 1971) ve temel beceriler (ör; gözlem yapma) ile üst düzey beceriler (ör; deney tasarlama) olmak üzere ikiye ayrılır. Ortaokul düzeyinde üst düzey becerilerin öğretilmesi tavsiye edildiği için (Sanderson & Kratochvil, 1971) ve bu çalışma 6. Sınıf (11-12 yaş grubu) öğrencileri ile yapıldığı için üst düzey bilimsel süreç becerileri çalışma kapsamında bilimsel okuryazarlığın bir parçası olan bilimsel süreç becerileri olarak düşünülmüştür.

Argümantasyon yöntemi bu noktada öğrencilerin bilimsel okuryazarlıklarını geliştirmek için kullanılabilir çünkü teorik olarak argümantasyon öğrencilerin içerik alan bilgisini (Jimenez-Aleixandre & Erduran, 2008), epistemolojik inançlarını (Sandoval & Millwood, 2008) ve bilimsel süreç becerilerini geliştirebilir. Bu noktada bazı temel kavramlar öne çıkmaktadır. Buna göre argümantasyon, bireylerin karşılıklı olarak birbirlerini ikna etmeye çalıştığı, birlikte hareket ettiği ve ortak bir noktada uzlaşıya vardıkları bir süreçtir (Cavagnetto, 2010) ve bu sürecin sonunda argümantasyonun ürünü olan argümanlar ortaya konur (Jimenez-Aleixandre & Erduran, 2008). Argümantasyonun yapıldığı sınıfların belli başlı bir takım özellikleri vardır. Buna göre, öğrenciler kendi öğrenmelerinden sorumludur ve süreç boyunca iddialarını kanıtla desteklemeleri gerekir. Öğretmen ise sınıfta otorite olarak yer almaz onun yerine süreç boyunca öğrencilere rehberlik eder. Öğretim programı ise öğrencilerin birer bilim insanı gibi çalışmasına olanak sağlamalı, gerçek hayattan problemlere yönelik olmalı ve uygulandığı zaman kişileri farklı sonuçlara yönlendirebilmelidir. Bu şekilde sınıfta tartışma ortamı gerçekleşir. Ölçme yöntemi ile ilgili olarak ise öğrencilerin sürece katılımı ve oluşturdukları argümanlar ile ilgili olarak tartışma öncesinde öğretmenler öğrencilere ölçütler sunmalı ve bu ölçütler göz önüne alınarak öğrenciler değerlendirilmelidir (Jimenez-Aleixandre, 2008).

Daha önce yapılan argümantasyon çalışmaları incelendiğinde araştırmacıların argümantasyonun ürünü olan argüman kalitesine odaklandıkları ve süreci ihmal ettikleri görülebilir (Sampson, Enderle, & Walker, 2012). Bu durumda öğrencilerin argümantasyon sürecine girerken neler yaptıkları ve nasıl tartıştıkları çok fazla bilinmemektedir. Örneğin, öğrencilerin hangi durumda argümantasyonda birlikte hareket ettikleri hangi durumda birbirlerinin iddialarını çürütmeye çalıştıkları bilinmemektedir (Sampson & Clark, 2011). Ayrıca, araştırmacılar daha önceki çalışmalarda argüman kalitesine odaklanmış olsalar bile Duschl (2007) daha önce yapılan argüman kalitesi ile ilgili çalışmaların sorunlu olduğunu iddia etmektedir. Çünkü söz konusu çalışmalar epistemik kriterleri karşılamamaktadır. Bu nedenle, Duschl (2007) argüman kalitesine odaklanmak isteyen araştırmacıların epistemik kriterleri karşılayan Walton'ın (1996) argüman şemalarını kullanmalarını önermektedir. Buna göre 25 farklı argüman şeması vardır ve bu şemalar günlük hayatta kullanılan argümanları işaret etmektedir. Argüman şemalarının kaynağında kesinlikle emin olamadığımız varsayımsal akıl yürütme becerileri vardır.

Bu çalışmada temel felsefe olarak etkileşimli yapılandırmacılık anlayışı benimsenmiştir. Etkileşimli yapılandırmacılık anlayışına göre doğada yer alan kurallar birşeyin doğru olup olmadığı konusunda bilim insanlarına yardımcı olur. İnsanların sosyal ortamda birbirleri ile yaptıkları tartışmalar sonucunda üretilen bilgi doğanın yasalarının süzgecinden geçirilir ve bilgi kanıtlara bağlı olarak kabul veya red edilir (Cavagnetto, Hand, & Norton-Meier, 2010).

Araştırmacıların argümantasyona karşı geliştirdikleri yönelimler üçe ayrılmaktadır bunlar derinlemesine yönelimler, sosyobilimsel yönelimler ve yapısal yönelimlerdir (Cavagnetto, 2010). Sosyobilimsel yönelimler ve yapısal yönelimler bilimsel sürecin içinde bulunduğu deney ve gözlemi ihmal edip bilimin ürününe odaklanırken, derinlemesine yönelimler hem bilimsel sürece hemde ürüne odaklanır (Cavagnetto, 2010), derinlemesine yönelimler bu bakımdan bilimsel okuryazarlığın vizyon-1 anlayışı ile uyumludur çünkü her iki kavram bilimin hem süreç hemde ürün boyutları ile ilgilidir. Bu nedenle bu çalışmada argümantasyon yönelimi olarak derinlemesine yönelimler tercih edilmiştir.

Bu çalışmada kavramsal çerçeve olarak derinlemesine yönelimler ve etkileşimli yapılandırmacılık felsefesi ile uyumlu olan Argüman Tabanlı Bilim Öğrenme (ATBÖ) yaklaşımı tercih edilmiştir. ATBÖ yaklaşımında öğretmen ve öğrencilerin yapmaları gerekenler temel hatlarıyla belirlenmiştir ve ATBÖ çalışmasının etkili bir şekilde uygulanması için öğretmen ve öğrencilerin bu beklentileri karşılaması gerekmektedir. Buna göre; öğretmen dersin başında öğrencilerin ön bilgisini öğrenmek için öğrencilere birer kavram haritası hazırlatır. Daha sonra öğretmen öğrencilerin araştırılabilir bir araştırma sorusu bulmasını ister. Öğrenciler araştırma sorusunu bulduktan sonra öğrencilerin araştırma sorularına nasıl cevap arayacakları sorulur. Bunun devamında öğrenciler araştırma sorularına cevap vermek için çalışmalarını yaparlar. Öğrenciler çalışmalarını tamamladıktan sonra bireysel argümanlarını oluştururlar, bu argümanlar araştırma sorularına cevap niteliğindedir. Daha sonra öğrenciler grup üyeleri ile bireysel argümanları hakkında tartışarak grup argümanını ortaya koyarlar. Grup argümanları oluşturulduktan sonra bu argümanlar sınıfın geri kalanına sunulur ve sınıf tartışması yapılır. Bu tartışmalar tamamlandıktan sonra tartışma sonuçların bilimsel bilgiler ile uyumlu olup olmadığı tartışılır bu noktada ders kitapları ve benzeri kaynaklar incelenir. Sürecin sonunda öğrenciler ders boyunca neler öğrendiklerini raporlandırır (Hand & Keys, 1999).

Bu çalışmanın ilgili alanyazın için önemli olduğu düşünülmektedir çünkü daha önce yapılan ATBÖ çalışmaları ATBÖ'nün alan bilgisine etkisine odaklanırken bu çalışma alan bilgisinin yanısıra epistemolojik inançların ve bilimsel süreç becerilerinin gelişimini incelemektedir. Ayrıca, daha önce yapılan argümantasyon çalışmalarının argüman kalitesini öğrenmek için kullandıkları yöntemler epistemik kriterleri karşılamamaktadır, Duschl (2007) bu noktada argüman kalitesinin argümantasyon şemaları (Walton, 1996) aracılığı ile analiz edilmesinin daha doğru olacağını düşünmektedir. Buna uygun şekilde mevcut çalışmada ortaokul öğrencilerinin argüman kaliteleri Walton'ın (1996) argümantasyon şemaları aracılığı ile analiz edilmiştir. Ayrıca ortaokul düzeyinde argümantasyon şemalarının bir analiz yöntemi olarak kullanıldığı bir ATBÖ çalışmasına rastlanılmamıştır. Bu özelliği nedeni ile mevcut çalışma alanda ilk olma özelliği taşımaktadır. Çalışmanın önemli olduğunu gösteren bir diğer durum çalışmanın argümantasyon sürecine odaklanmasıdır. Kim ve Song (2006) daha önce yapılan argümantasyon çalışmalarının çoğunlukla argümantasyonun sonucu olan argümana odaklandığını göstermektedir fakat argümantasyon sürecine odaklanan çok fazla çalışma bulunmamaktadır. Bu nedenle argümanlar ortaya konmadan önce oluşan süreçler hakkında çok fazla bilgi bulunmamaktadır. Söz konusu çalışma argümantasyon sürecini araştırarak, bu süreç hakkında bilgi sunmaktadır. Bu çalışma ayrıca içinde yapılmış olduğu bağlamın argümantasyona ne derece uyumlu olduğu hakkında bilgi vermektedir. Alanyazından elde edilen bulgulara bağlı olarak Doğu kültüründe ve okullarında argümantasyonın çok fazla yapılmadığını fakat Batı kültüründe ve okullarında argümantasyonun daha fazla yapıldığı söylenebilir (Lin & Mintzes, 2010). Türkiye ise gerek coğrafya gerekse kültür olarak bu iki kültürün etkisi altındadır. Bu nedenle mevcut çalışma bu iki kültürün arasında kalan Türkiye bağlamının argümantasyona ne derece uyumlu olduğu konusunda bilgi vermektedir. Ayrıca bu çalışmada ATBÖ yaklaşımının öğrencilerin fen okuryazarlığını geliştireceği hipotezi ortaya atılmıştır. Eğer beklendiği gibi bütün fen okuryazarlığı boyutları ATBÖ'ye bağlı olarak gelişirse, farklı örneklemlerle yapılacak çalışmalar ile ATBÖ'nün fen okuryazarlığı üzerindeki etkisinin hangi koşullarda genellenebileceği ortaya konabilir. Benzer şekilde, eğer bu çalışma ile öğrencilerin fen okuryazarlık boyut veya boyutları gelişmezse neden bu boyutların gelişmediği ile ilgili olarak araştırmacılar yeni araştırmalar yapabilirler. Son olarak bu çalışmada öğrencilerin en çok kullandığı ve en az kullandığı argümantasyon şemaları ile öğrencilerin en çok kullandıkları ve en az kullandıkları argümantasyon sürecine katılım boyutları öğretmenleri hazırlayacakları ders planları konusunda bilgilendirebilir. Çalışmada geliştirilen ve kullanılan etkinlikler ayrıca öğretmenler tarafından kullanılabilir ve çalışmada karşılaşılan zorluklar öğretmenlerin ATBÖ yaklaşımını uygularken nelere dikkat etmesi konusunda bilgi verebilir.

**Araştırma Soruları:**

Bu çalışma beş soruya cevap aramaktadır, bu sorular:

1. ATBÖ yaklaşımı ile öğrenim gören 6. sınıf öğrencilerinin uygulamadan önceki ve uygulamadan sonraki içerik alan bilgileri arasında anlamlı bir fark var mıdır?

2. ATBÖ yaklaşımı ile öğrenim gören 6. sınıf öğrencilerinin uygulamadan önceki ve uygulamadan sonraki epistemolojik inançları arasında anlamlı bir fark var mıdır?

3. ATBÖ yaklaşımı ile öğrenim gören 6. sınıf öğrencilerinin uygulamadan önceki ve uygulamadan sonraki bilimsel süreç becerileri arasında anlamlı bir fark var mıdır?

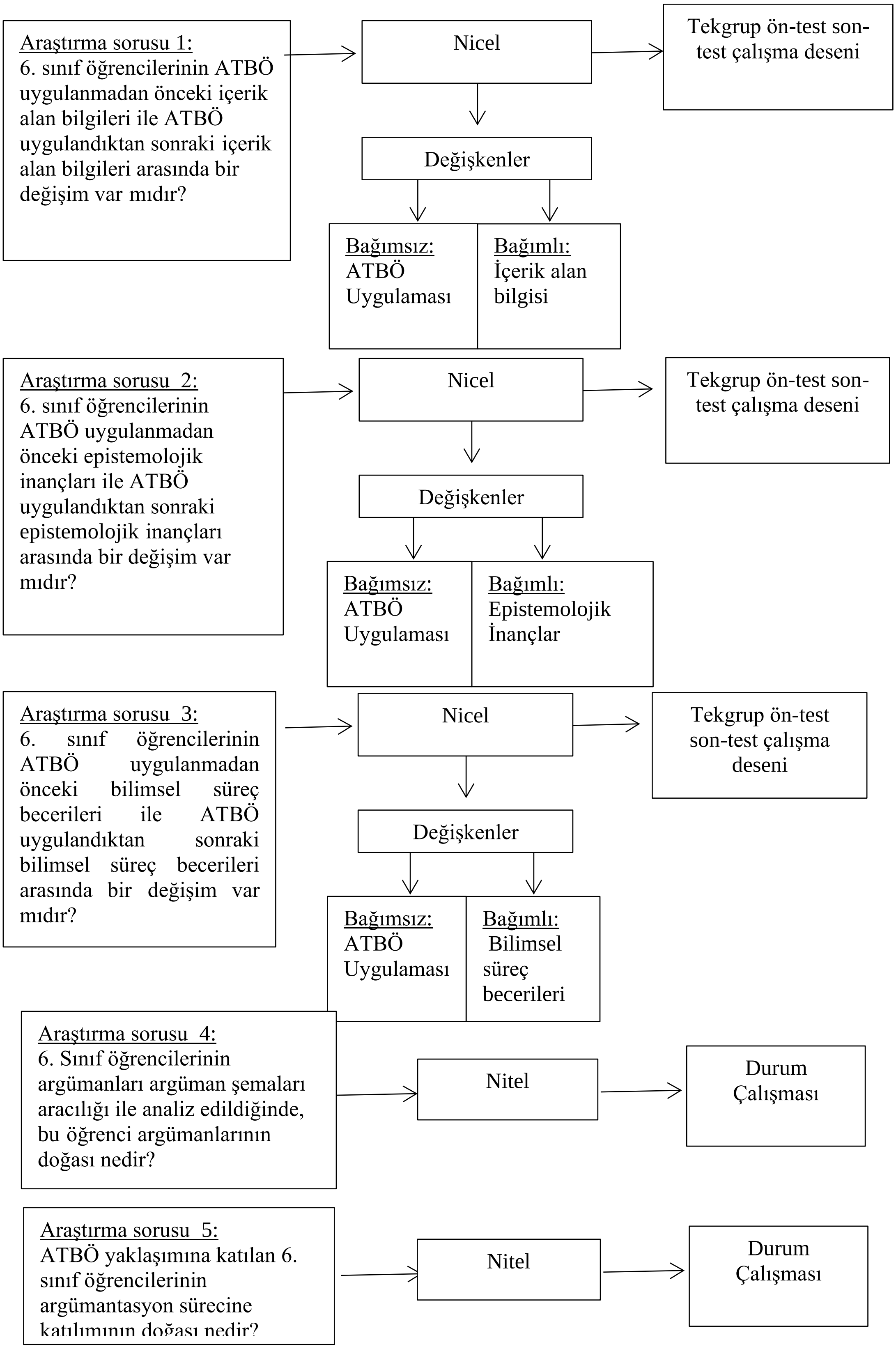
4. 6. Sınıf öğrencilerinin argümanları argüman şemaları aracılığı ile analiz edildiğinde, bu öğrenci argümanlarının doğası nedir?

5. ATBÖ yaklaşımına katılan 6. sınıf öğrencilerinin argümantasyon sürecine katılımının doğası nedir?

## YÖNTEM

### Çalışma Deseni

Bu çalışmada yer alan ilk üç araştırma sorusu nicel yöntemlerle son iki araştırma sorusu ise nitel yöntemlerle araştırılmıştır. Dolayısı ile bu çalışmanın çalışma deseni hem nicel hem nitel araştırmayı içerir. Ancak, bu iki çalışma deseninden elde edilen veriler birbirlerini desteklemek için kullanılmamıştır. Bu yüzden bu çalışma karma çalışma deseni olarak değil, nitel ve nicel araştırmanın birlikte kullanımı olarak isimlendirilmiştir (Creswell, 2006, p.83). Araştırmanın nicel kısmında tek grup ön-test son-test deseni kullanılmıştır. Burada Ankara'da merkez bir ilçede yer alan bir devlet okulunda öğrenim gören dört adet 6. sınıfa ATBÖ uygulaması verilmiştir ve bu nedenle grupların tamamı deney grubu olmuştur. Çalışmanın nitel kısmı ise durum çalışmasıdır. Buna göre nicel araştırmaya katılan dört sınıftan iki tanesi nitel araştırma için seçilmiştir. Çalışmada yer alan ABI uygulaması bu çalışmadaki durumu (örnek olay) temsil etmektedir. Çalışmaya katılan iki sınıf ise bu durumun alt birimlerini oluşturmaktadır. Aşağıda yer alan şekil araştırma sorularını, ilgili çalışma desenlerini ve araştırma türlerini göstermektedir (Şekil 1).




*Şekil 1.* Çalışma Deseni ve Türleri

**Katılımcılar**

Çalışmanın nicel kısmının çalışma evreni Ankara ilinde devlet okullarında öğrenim gören 6. sınıf öğrencileridir. Çalışmanın ulaşılabilir evreni ise Çankaya ilçesinde devlet okullarında öğrenim gören 6. sınıf öğrencileridir. Örneklem çeşidi olarak kolay ulaşılabilir durum örneklemesine gidilmiş ve buna bağlı olarak bir devlet okulunda pilot ve ana çalışma yürütülmüştür. Aynı öğretmenin öğretim verdiği dört sınıfta yer alan toplam 71 öğrenci (36 erkek-35 kız) çalışmaya katılmıştır. Çalışmanın nitel kısmı içinse ATBÖ eğitiminin verildiği dört sınıftan ikisi rastgele seçilmiş ve seçilen iki sınıfta nitel çalışma yürütülmüştür. Çalışmanın nitel kısmına bir şubeden 17, diğer şubeden 18 öğrenci olmak üzere toplam 35 öğrenci katılmıştır. Araştırma okulun fen laboratuvarında gerçekleştirilmiş ve öğrenciler gruplar halinde dört ayrı masada çalışmalarını yapmışlardır. Her bir grup 4-5 öğrenciden oluşmuştur.

**Çalışma Öncesi Yapılanlar**

Çalışma başlamadan önce ATBÖ yaklaşımı ile uyumlu olduğu düşünülen Madde ve Isı ünitesi ile Elektrik ünitesi seçilmiştir çünkü bu iki ünitenin kazanımları hem deney hem tartışma ile ilgilidir ve bu durum ATBÖ uygulaması için uygundur. Daha sonra araştırmacı ATBÖ yaklaşımının doğru bir şekilde uygulanması için ilgili konularda ders planları hazırlamıştır. Her bir ders planı hazırlanırken ATBÖ safhaları göz önünde bulundurulmuştur (Hand & Keys, 1999), bu durumun çalışmanın geçerliğini artırdığı düşünülmektedir. Alan uzmanlarından alınan dönütlerle ders planlarında düzenlemeler yapılmıştır. Hazırlanan ders planlarının ana düşünceleri ısı iletimi, ısı yalıtımı, termal yalıtım malzemeleri, yakıt türleri, elektrik iletimi ve ampül parlaklığını etkileyen etmenlerdir. Buna göre hem pilot çalışma hemde ana çalışma 6 hafta sürmüş ve bu 6 haftanın ilk 4 haftasında madde ve ısı ünitesi işlenirken son 2 haftasında elektrik ünitesi işlenmiştir. Öğrenciler çalışmanın ilk 2 haftası ve son 2 haftasında deneyler yaparak doğrudan veri toplamışlar, buna karşın katılımcılar çalışmanın 3. ve 4. haftalarında araştırmacının hazırlamış olduğu kanıt kağıtlarını kullanarak dolaylı yoldan veri toplamışlardır.

Çalışma başlamadan önce ODTÜ Sosyal Bilimler Enstitüsü Etik Kurulu ve İl Milli Eğitim Müdürlüğü Etik Kurulu'ndan hem pilot hem ana çalışma için etik izinleri ve uygulama izinleri alınmıştır. Daha sonra çalışmanın yapılacağı okuldan izinler alınıp çalışmayı yürütecek öğretmen ile tanışılmıştır. ATBÖ eğitiminin verimli bir şekilde verilmesi için çalışma öncesinde çalışmaya katılan öğretmene 6 ayrı oturumda bilim, bilimsel süreç becerileri, bilimin doğası, argümantasyon, ATBÖ ve sınıfta argümantasyon dersinin değerlendirilmesi ile ilgili bilgiler verilmiştir. Gerek pilot çalışmada gerekse ana çalışmada araştırmacı ve uygulayıcı öğretmen arasında fikir alışverişi yapılarak ATBÖ'nün doğru bir şekilde uygulanmasına çalışılmıştır.

Pilot çalışma 2017-2018 Bahar döneminde yapılmıştır. Pilot çalışmada her bir uygulama sonunda araştırmacı ATBÖ uygulamasının geliştirilmesi için notlar almıştır. Alınan bu notlar ile bir sonraki sene yapılan ana çalışmanın daha verimli yapılması amaçlanmıştır. Örneğin, pilot çalışmanın son haftasında öğrenciler telin direncini etkileyen etmenleri incelerken telin boyu ve direnci arasındaki ilişkiyi gösteren deneyi yapamamışlardır, bu nedenle ana çalışmada söz konusu deney aynı ilişkiyi gösteren bir simülasyon ile değiştirilmiştir.

**Veri Toplama Araçları**

Bu çalışmada 4 farklı araç kullanılarak veri toplanmıştır. Katılımcıların içerik alan bilgilerini ölçmek amacıyla araştırmacı tarafından geliştirilen her biri 20 adet çoktan seçmeli madde ve ısı testi ile elektrik testi kullanılmıştır. Testlerin oluşturulması sırasında geçerliğin sağlanması için alan uzmanlarından görüşler alınmıştır. Pilot çalışmada madde ve ısı testi 157 öğrenciye, elektrik testi 195 öğrenciye uygulanmıştır. Her bir madde için madde zorluk endeksi ve madde ayırıcılığı gücü hesaplanmıştır ve maddelerin tamamının testlerde kalması uygun görülmüştür. KR-20 değerleri kullanılarak test sonuçlarının iç tutarlılık güvenilirliği hesaplanmıştır. Buna göre madde ve ısı test sonucu güvenilirliği ön-testte .64, son-testte .76, elektrik test sonucu güvenilirliği ön-testte ve son-testte .76 olarak hesaplanmıştır. Sonuç olarak geliştirilen içerik alan bilgisi testlerinin güvenilir ve geçerli olduğu söylenebilir.

Katılımcıların epistemolojik inançlarını ölçmek için Conley vd. (2004) tarafından geliştirilen epistemolojik inanç ölçeği (EBQ) kullanılmıştır. Ölçek 26 maddedir ve bilimsel bilginin kaynağı, kesinlği, gerekeçelendirilmes ile gelişimi boyutlarını içermektedir. Bu ölçek Özkan (2008) tarafından Türkçe'ye çevrilmiş ve adapte edilmiştir. Özkan'ın bulgularına uygun şekilde bilimsel bilginin kaynağı ve kesinliği boyutları bu çalışmada birleştirilerek tek faktör olarak ele alınmıştır. Bilimsel bilginin gerekçelendirilmesi ve gelişimi boyutları ayrı birer faktör olarak korunmuştur. Sonuç olarak katılımcıların epistemolojik inançları 3 ayrı boyut olarak araştırılmıştır. Çalışmanın güvenilirliği ile ilgili olarak Cornbach alfa değeri çalışma öncesinde 0.81, çalışma sonrasında ise 0.91 olarak bulunmuştur.

Burn vd. (1985) tarafından geliştirilen bilimsel süreç becerileri testi (TIPSII) katılımcıların bilimsel süreç becerilerini ölçmek için kullanılmıştır. Bu test çoktan seçmeli 36 soruyu barındırmaktadır ve özel olarak herhangi bir fen konusu ile ilişkili değildir. Test Geban, Akar ve Özkan (1992) tarafından Türkçe'ye uyarlanmıştır. Can (2008) ise ortaokul seviyesinde ölçeği uygulamış ve ölçekten ayırt ediciliği düşük olan 10 maddeyi çıkartmıştır. 26 maddelik bu ölçek mevcut çalışmada da kullanılmıştır. Bilimsel süreç becerileri alt boyutları güvenilirlik değerleri çok düşük olduğu için ölçek tek boyutlu yorumlanmıştır. Bu çalışmada ölçeğin KR-20 güvenilirlik değeri ön-testte .83, son-testte .86 olarak hesaplanmıştır.

Çalışmanın son iki sorusu olan nitel araştırma sorularına cevap aramak için sınıf içi gözlemler yapılmıştır. Bu gözlemler sırasında video kaydı yapılmıştır. Video kayıtları araştırmacı tarafından izlenmiş ve tüm sınıf argümantasyonun yapıldığı bölümler her bir ders için tanımlanarak transkript edilmiştir.

**Ana Çalışma**

Ana çalışma 2018-2019 yılı bahar döneminde yapılmıştır. Çalışma başlamadan önce çalışmaya katılan öğrencilere eğitim verilmiştir. Bu eğitimde bilimsel tartışma, merak, farklı düşüncelerin önemi, bilim insanı, bilim insanı ve öğrenciler arasındaki benzerlikler, bilim insanının nasıl çalışma yaptığı, bilimsel süreç becerileri, örnek ATBÖ dersi ve ATBÖ uygulaması öncesi uyulacak kurallar üzerinde durulmuştur.

Daha sonra örnek bir uygulama ile öğrencilerin argümantasyona katılması teşvik edilmiş ve ders sonunda buna benzer etkinlikler yapılacağı bilgisi öğrencilerle paylaşılmıştır. Pilot çalışmadan alınan dönütler doğrultusunda ders planlarında düzenlemeye gidilmiş ve düzenlenen ders planlarının uygulanması ile ana çalışma yapılmıştır. Örneğin birinci haftada önbilgilerin yoklanması amacıyla birinci safhada öğrencilerin pikniğe gittiklerinde yapacakları yemeği hızlı bir şekilde pişirmeleri için hangi kabı kullanmaları gerektiği sorulmuştur. Sınıf içerisinde yapılan bu ön tartışma ile öğrencilerin maddelerin ısı iletimi ile ilgili ön bilgileri alınmıştır. Daha sonra ikinci safhada öğrencilerden ısı iletimi ile ilgili araştırma soruları yazmaları ve deney tasarlamaları istenmiştir. Yazılan araştırma soruları araştırılabilirlik ve konu ile ilgisi bakımından sınıf üyeleri tarafından tartışılmıştır. Daha sonra ilgili ölçütleri karşılayan araştırma soruları deneyler aracılığı ile üçüncü aşamada test edilmiştir. Öğrenciler gruplar halinde yaptıkları deney sonuçlarını dördüncü safhada bireysel argümanlarını oluşturmak üzere kullanmışlardır. Beşinci safhada öğrencilerden bireysel argümanlarını grup arkadaşları ile tartışmaları istenmiş ve bu tartışmaların sonucunda grubun ortak argümanı ortaya konmuştur. Altıncı aşamada gruplar deneylerini ve grup iddialarını sınıfın geri kalanına sunmuşlardır. Bu aşamada tüm sınıfın dahil olduğu argümantasyon süreci başlamıştır. Yapmış oldukları deneylerde bazı gruplar metal kaşığın bazı gruplar ise plastik kaşığın daha iyi ısı iletkeni olduğunu bulmuşlardır. Gruplar kanıtları ve veri toplama süreçlerini inceledikten ve tartıştıktan sonra ortak noktalarda buluşmuşlardır. Daha sonra gruplar neden bazı maddelerin daha iyi ısı iletkeni olduğu konusunda teorik sorular sormuş ve bu sorulara cevaplar aramışlardır. Teorik soruların cevaplanması için öğretmende tartışma sürecine dahil olmuş ve bu tartışmanın sonucunda bazı maddeleri oluşturan taneciklerin diğer maddeleri oluşturan taneciklere göre daha sık ve düzenli olmasından dolayı sık ve düzenli taneciklere sahip  maddelerin daha iyi iletken olduğu bilgisine ulaşılmıştır. Daha sonra öğrenciler maddeleri ısıyı iyi ileten maddeler ve ısıyı yalıtan maddeler olarak sınıflandırmışlar ve ilgili dersin kazanımına ulaşmışlardır. Sürecin sonunda öğrenciler öğrenci raporlarına bu dersten neler öğrendiklerini yazmışlardır. Her bir haftanın etkinlği toplam 4 ders saati sürmüştür. Birinci haftaya benzer şekilde diğer haftaların etkinliklerinin yapılması ile 6 hafta sonunda ana çalışma tamamlanmıştır. Çalışma boyunca ATBÖ'nün verimli bir şekilde yapılıp yapılmadığını test etmek amacıyla araştırmacı Sawada, Piburn, Falconer, Turley, Benford ve Bloom (2000) tarafından geliştirilen Reform Edilmiş Öğretim Gözlem Protokolü (RTOP)'u kullanmıştır. Buna göre gözlem protokolünde bulunan maddelerden ATBÖ ile doğrudan ilgili olan 13 maddenin kontrol listesi her derste araştırmacı tarafından doldurulmuş ve ilgili dönütler uygulayıcı öğretmenle ders sonunda paylaşılmıştır.

**Veri Analizi**

Bu çalışmada ilk üç araştırma sorusuna yanıt aramak için nicel veri analizi, son iki araştırmasına yanıt aramak içinse nitel veri analizi yapılmıştır. Birinci araştırma sorusu ATBÖ yönteminin içerik alan bilgisi üzerine etkisi ile ilgili olduğu için grup içerisinde tekrar eden çoklu varyans analizi uygulanmıştır. Bu analizde bağımlı değişkenler öğrencilerin madde ve ısı konusundaki ve elektrik konusundaki içerik alan bilgileridir. Zaman ise ATBÖ uygulamasını temsil eden bağımsız değişken olarak belirlenmiştir. Benzer şekilde ikinci araştırma sorusunda ATBÖ yaklaşımının öğrencilerin epistemolojik inançlarına etkisi üzerinde durulmuştur. Bu soruya yanıt vermek için bir önceki araştırma sorusunda olduğu gibi grup içerisinde tekrar eden çoklu varyans analizi kullanılmıştır. Bu analizde zaman yine bağımsız değişken olarak kullanılırken, bağımlı değişkenler epistemolojik inancın boyutları olan bilimsel bilginin gerekçelendirmesi, bilimsel bilginin gelişimi ve bilimsel bilginin kaynağı ile kesinliğidir. Üçüncü araştırma sorusunda ATBÖ yaklaşımının öğrencilerin bilimsel süreç becerilerine etkisi sorulmuştur. Bilimsel süreç becerileri tek bir boyut olarak ele alınmış ve bağımlı değişken olarak düşünülmüştür. Bağımsız değişken ise önceki iki analizde olduğu gibi zaman olarak belirlenmiştir. Bu nedenle, ATBÖ'nün bilimsel süreç becerilerine etkisini analiz etmek için bağımlı örneklem t-testi kullanılmıştır.

Çalışmanın son iki araştrma sorusu için nitel analiz yöntemlerinden sürekli karşılaştırmalı analiz kullanılmıştır. Dördüncü araştırma sorusuna cevap verebilmek için Walton'ın (1996) argümantasyon şemaları (ör; işaret yoluyla oluşturulan argüman) kullanılarak tümden gelimsel kodlama yapılmıştır. Argümantasyon şemaları ortaya konduktan sonra her bir sınıf için her hafta toplam argümantasyon şeması çıkartılmış ve sütun grafiği ile sunulmuştur. Argüman şemalarının birbiri ile kıyaslanması içinse her hafta için yüzdelik dilimler oluşturulmuş ve daire grafiği ile sunulmuştur. Sonuncu araştırma sorusu öğrencilerin argümantasyona sürecine katılımlarını ortaya koymayı amaçlıyordu. Bu nedenle Sampson ve Clark'ın (2011) ortaya koyduğu argümantasyon katılım boyutları kullanılarak (ör; bilgi arama) analizlerde tümden gelimsel kodlama yapılmıştır. Argümantasyona katılım boyutları oluşturulduktan sonra her sınıf için haftalık toplam katılım sıklığı oluşturulmuş ve bu toplam katılım sayıları sütun grafiği kullanılarak sunulmuştur. Benzer şekilde haftalık olarak ortaya konan argümantasyona katılım boyutları her hafta birbiri ile kıyaslanarak yüzdelik dilimleri hesaplanmış ve bu yüzdelik dilimler daire grafiği ile sunulmuştur. Gerek dördüncü araştırma sorusunun veri analizinde, gerekse sonuncu sorunun veri analizinde ilgili grafikler sunulduktan sonra her iki sınıfın sonuçları birbirleri ile karşılaştırılmıştır. Sınıfların sonuçları arasındaki benzerlikler ve farklılıklara bağlı olarak 6. sınıf öğrencilerinin ATBÖ dersinde ortaya koydukları argümantasyon şemaları ve öğrencilerin argümantasyon sürecine katılımları ile ilgili iddialar ortaya atılmıştır. Bu iddialar öğrencilerin argümantasyon şemalarının ve argümantasyon sürecine katılımının doğasını yansıtmaktadır.

## SONUÇLAR ve TARTIŞMA

Bu çalışmada ilk olarak ATBÖ uygulamasının 6. sınıf öğrencilerinin içerik alan bilgisi üzerine etkisi incelenmiştir. Grup içi tekrarlanan çoklu varyans analizi ile bu soruya yanıt aranmadan önce, bu analiz yönteminin sayıltıları kontrol edilmiş ve analizi yapmaya engel olacak bir durumla karşılaşılmamıştır. Grup içi tekrarlanan çoklu varyans analizi sonuçlarına göre katılımcıların içerik alan bilgisi ATBÖ yöntemine bağlı olarak istatistiksel olarak anlamlı ölçüde gelişmiştir, $F(2,67)= 20.55$, $p= 0.00$; Wilks' Lambda= 0.62. Çoklu varyans analizinden sonra hangi ünitede katılımcıların içerik alan bilgisinin geliştiğini saptamak amacıyla bağımlı örneklem t-testi yapılmıştır. Bağımlı örneklem t-testi sonuçlarına göre katılımcıların madde ve ısı ünitesi içerik alan bilgileri ön-testten ($M$=9.64, $SD$=3.37) son-teste ($M$=11.54, $SD$=4.13) istatistiksel olarak anlamlı bir şekilde artmıştır, $t(68)=4.48$, $p<0.025$ (iki uçlu test). Madde ve ısı konusu için etki büyüklüğü 0.23 olarak hesaplanmıştır. Benzer şekilde katılımcıların elektrik ünitesi içerik alan bilgisi değişimini anlamak amacıyla bağımlı örneklem t-testi yapılmış ve ön-testten (*M*=9.52, *SD*=4.25) son-teste (*M*=11.99, *SD*=4.28) istatistiksel olarak katılımcıların elektrik ünitesi içerik alan bilgisinde anlamlı bir artış olmuştur $t$ (68) =5.47, $p<0.025$ (iki uçlu test). Elektrik konusu için etki büyüklüğü 0.31 olarak hesaplanmıştır.

Bu çalışmada kontrol grubu kullanılmadığı için ATBÖ uygulamasının içerik alan bilgisini geliştirdiği kesin olarak iddia edilemez fakat bazı göstergeler ATBÖ uygulamasının içerik bilgisini geliştirdiği fikrini desteklemektedir. Örneğin; etki büyüklüğünün ikinci ünitede birinci üniteye göre daha yüksek olması ATBÖ yönteminin içerik alan bilgisini geliştirdiği düşüncesini destekler. ATBÖ'nün biliminsel sürece vurgu yapması ve bu sürece katılan öğrencilerin de bilgilerini geliştirmesi ATBÖ'nün içerik alan bilgisini desteklediğini gösteren bir diğer unsurdur (Kingir vd., 2013). ATBÖ yaklaşımının ayrıca kavramsal değişim yaklaşımı ile uyumlu olması, öğrencilerin ön bilgilerindeki hataların farkına varması ve doğru bilgiyi edinmeleri ATBÖ aracılığı ile içerik alan bilgilerini geliştirmelerine katkıda bulunmuş olabilir (Kingir vd., 2013). Ayrıca öğrencilerin ATBÖ süreci boyunca kullandıkları yazılı raporlar öğrenmelerini kolaylaştırmış olabilir çünkü bu yazılı raporlar argüman üretilmesi ve fikirlerin paylaşılması gibi konularda öğrencilerin düşünmesini sağlayan sorular içermektedir (Akkus vd., 2007; Chen vd., 2016; Cronje vd., 2013; Hand vd., 2004). ATBÖ'nün içerik alan bilgisini artırmasını sağlayan bir diğer etmen, ATBÖ aracılığı ile öğrencilerin ön tartışmalara, grup içi tartışmalara ve sınıf tartışmasına katılması olabilir. Bu şekilde öğrenciler birbirlerinden öğrenme ve birlikte öğrenme yoluna giderek içerik bilgilerini geliştirmiş olabilir (Hand vd., 2004; Heng vd., 2015; Shamuganathan & Karpudewan, 2017). Benzer şekilde, bu çalışmada ATBÖ'nün bütün safhalarının dikkatlice uygulanması içerik bilgisini geliştirmiş olabilir çünkü daha önce yapılan bazı çalışmalarda ATBÖ'nün bütün safhalarının kullanılmaması sonucunda uygulamanın istendiği kadar verimli olmadığı raporlanmıştır (Cronje vd., 2013; Erkol vd., 2007).

İkinci araştırma sorusunda ATBÖ yaklaşımının öğrencilerin epistemolojik inançlarına etkisi araştırılmış ve veri analizi için grup içi tekrarlanan çoklu varyans analizi kullanılmıştır. Çoklu varyans analizi yapılmadan önce sayıltıları test edilmiştir. Grup içi tekrarlanan çoklu varyans analizi sonuçlarına göre ATBÖ yaklaşımının katılımcıların epistemolojik inançları üzerinde istatistiksel olarak anlamlı bir etkisi bulunmamıştır $F(3,62)= 2.24$, $p= 0.09$; Wilks' Lambda= 0.90.

Araştırmada öğrencilerin epistemolojik inanç ön-test sonuçları ortalaması 5 üzerinden yaklaşık olarak 4 seviyesindeydi. Bu nedenle halihazırda yüksek olan epistemolojik inançlar son-testte daha fazla gelişmemiş olabilir. Ayrıca katılımcılara doğrudan açık bir şekilde epistemolojik inanç eğitimi verilmediği için de katılımcıların epistemolojik inançları gelişmemiş olabilir (McDonald & McRobbie, 2012; McDonald, 2010). Ayrıca çalışma süresinin görece olarak kısa olması (6 hafta), ATBÖ'nün verimliliğini azaltmış olabilir ve bu nedenle katılımcıların epistemolojik inançları gelişmemiş olabilir. Ayrıca nicel sonuçlar öğrencilerin çalışma öncesinde ve sonrasında gelişmiş epistemolojik inançlara sahip olduğunu göstersede gözlem sonuçları öğrencilerin epistemolojik inançlarının gelişmemiş olduğunu işaret etmektedir. Öğrencilerin daha önce almış oldukları geleneksel eğitimdeki içerik alan bilgisi vurgusu, kendilerine verilen kanıt kağıtlarına karşı eleştirel bir tutum sergilememeleri, katılımcıların epistemolojik inanç gelişimi için yeterli gelişimsel olgunlukta olmamaları, tartışmaya daha az katılan sınıfın deney yapılan haftalarda pasif kalması öğrencilerin gelişmemiş epistemolojik inanca sahip olduğu düşüncesini desteklemektedir.

Üçüncü araştırma sorusunda ATBÖ'nün bilimsel süreç becerileri üzerine etkisi araştırılmıştır. Bilimsel süreç becerileri tek bir boyut olarak ele alındığı için çoklu varyans analizi yerine bağımlı örneklem t-testi yapılmıştır. Bağımlı örneklem t-testi sonuçlarına göre katılımcıların bilimsel süreç becerilerinde ön-testten ($M$=14.70, $SD$=5.48) son-teste ($M$=14.33, $SD$=5.98) istatistiksel olarak anlamlı bir değişim olmamıştır $t\ (62) = -0.59$, $p>0.05$ (iki uçlu test). Bu nedenle ATBÖ'nün katılımcıların bilimsel süreç becerileri üzerinde anlamlı bir etkisi olmadığı söylenebilir.

Epistemolojik inançta olduğu gibi bilimsel süreç becerilerinin de gelişmemesinin öncelikli nedeni çalışmanın görece olarak kısa süreli olması olabilir. Benzer şekilde, Yıldırım (2012) da kısa süreli olan rehberli araştırma-sorgulamaya dayalı öğrenmenin öğrencilerin bilimsel süreç becerisini geliştirmede yetersiz olduğu sonucuna varmıştır. Yine benzer şekilde öğrencilere uygulama sırasında doğrudan bilimsel süreç becerileri eğitimi verilmediği için ve öğrenciler katıldıkları etkinliklerde açık bir şekilde bilimsel süreç becerilerini ifade etmedikleri için bu çalışmada öğrencilerin bilimsel süreç becerileri gelişmemiş olabilir.

Dördüncü araştırma sorusunda 6. sınıf öğrencilerinin argümantasyon şemasına göre incelendiğinde ortaya koydukları argümanların doğası araştırılmıştır. Buna göre her iki sınıfın argüman şemaları Walton'ın (1996) argümantasyon şemalarına göre analiz edilmiş ve sınıflar arası benzerliklere ve farklılıklara dayanarak beş iddia ortaya atılmıştır. Sınıflar arası benzerliklere bağlı olarak ortaya atılan iddialar şu şekildedir:

İddia 1. Çalışma süresince öğrencilerin kullanmış oldukları argümantasyon şeması çeşidi artmıştır.

Birinci iddiada yer alan zamanla katılımcıların daha fazla argümantasyon şeması kullanması öğrencilerin zamanla daha fazla akıl yürüttüğünü ve muhakeme yaptığını göstermektedir (Konstantinidou & Macagno, 2013). Duschl'un (2007) yapmış olduğu ATBÖ'nün uygulanmadığı bir başka çalışmada ortaokul öğrencileri yalnızca 9 farklı argümantasyon şeması kullanmıştır. Ancak bu çalışmada ATBÖ uygulanması ile öğrencilerin kullanmış oldukları şema çeşidi sayısının artmış olduğu ve toplamda 17 farklı şema kullandıkları gözlemlenmiştir. Bu nedenle ATBÖ yaklaşımı öğrencilerin farklı argümantasyon şemaları kullanmasını teşvik ediyor olabilir.

İddia 2. Öğrenciler uzman görüşü yoluyla ortaya konan, işaret yoluyla ortaya konan, eksik bilgi sonucu ortaya konan, benzerlik yoluyla ortaya konan ve hipotez yoluyla ortaya konan argümantasyon şemalarını diğer argümantasyon şemalarından daha çok tercih etmektedir.

Öğrencilerin daha fazla kullanmış oldukları bu beş farklı argümantasyon şeması ortaokul düzeyinde sınıf tartışması sırasında kullanılan bilimsel argümantasyon şemaları olabilir. Örneğin; öğrenciler kendilerine kanıt kağıdı verildiğinde uzman görüşü yoluyla ortaya konan argüman şemasını kullanmaktadır. Benzer şekilde gözlemledikleri bir durumu açıklamak için çıkarımlar yaparak sorularına cevap aramakta bu nedenle Duschl'un (2007) da belirttiği gibi işaret yoluyla ortaya konan argümantasyon şemasını kullanmaktadırlar. Öğrenciler diğer grupların bulgularının ve sonuçlarının tutarsız olduğunu düşünürse yada diğer grupların bilgilerinin eksik olduğunu düşünürse eksik bilgi sonucu ortaya çıkan argüman şemasını kullanmaktadır. Ayrıca, öğrenciler yapmış oldukları deneyleri açıklarken deney ve kontrol gruplarını karşılaştırmak amacıyla sıklıkla benzerlik yoluyla ortaya konan argümantasyon şemasını kullanmaktadır. Son olarak deney öncesi yaptıkları hipotezleri ve değişkenlerini açıklarken öğrenciler sık sık hipotez yoluyla ortaya konan argümantasyon şemasını kullanmışlardır.

İddia 3. Öğrencilerin kullanmış oldukları veri türü (ör; doğrudan, dolaylı) öğrencilerin kullanmış oldukları argümantasyon şeması tercihlerini etkileyebilir.

Bu çalışmada öğrencilerin deney yaparak kendi verilerini topladıkları haftada hipotez yoluyla ortaya konan argümantasyon şemasını, deney yapmadıkları haftalarda ise uzman görüşü aracılığıyla ortaya konan argümantasyon şemasını kullanmayı tercih etmişlerdir. Bu nedenle öğrencilerin doğrudan veya dolaylı olarak elde ettikleri veriler kullandıkları argüman şemasını etkilemiştir. Benzer şekilde, Kind vd. (2011) öğrencilerin belli bir yönelimleri olmadığını ve deney yapmaları gerektiğinde deneyi yapmaya odaklandıklarını, deney olmadığı zaman ise bilimin ürünü olan içerik bilgisine odaklandığını rapor etmişlerdir. Yine aynı şekilde Özdem vd. (2013) öğrencilerin yapmaları gereken etkinliğe bağlı olarak ilişki yoluyla ortaya konan argümantasyon şemalarını veya işaret yoluyla ortaya konan argümantasyon şemalarını kullandıklarını rapor etmişlerdir.

Sınıflar arası farklılıklara bağlı olarak argümantasyon şemaları ile ilgili ortaya atılan iddialar ise şu şekildedir:

İddia 1. Argümantasyon sürecine daha fazla dahil olan sınıf diğer sınıfa göre daha fazla argümantasyon şeması kullanmaktadır.

Bu çalışmada sınıflardan birisi diğerinden daha fazla argümantasyon sürecine dahil olmuştur ve daha fazla sürece katılan sınıfın daha fazla argümantasyon şeması çeşidi kullandığı gözlemlenmiştir. Buna göre daha fazla argümantasyon sürecine girmiş olan grup daha fazla konu üzerinde konuşup, daha fazla fikir öne sürmüş olabilir (Sampson & Clark, 2011). Konu üzerine daha fazla düşünülmesi de daha fazla muhakeme yapıldığını ve bunun sonucunda aktif olan sınıfın daha fazla argümantasyon şeması kullandığını gösteriyor olabilir.

İddia 2. Daha aktif olarak argümantasyon sürecine katılan sınıf eksik bilgi sonucu ortaya konan argüman şemasını en çok kullanırken, daha az aktif olarak argümantasyon sürecine katılan sınıf uzman görüşü yoluyla ortaya konan argüman şemasını en çok kullanmaktadır.

Daha aktif grubun en çok eksik bilgi nedeniyle argümantasyon şeması kullanması ve daha az aktif sınıfın en çok uzman görüşü yoluyla argümantasyon şeması kullanması bu iki sınıftan neden birinin diğerinden daha fazla argümantasyon sürecine katıldığını açıklıyor olabilir. Lederman ve Lederman (2012) bilimin soru sorarak başladığını belirtmektedir ve aktif olan sınıf sürekli sorular sorarak bilimsel süreci başlatmış ve süreç boyunca bilgiyi aramıştır ve bu süreçte argümantasyona katılmışlardır. Bir diğer ifade ile aktif olan sınıf öğrencileri bilim insanı gibi çalışmışlardır. Öte yandan daha az aktif olan sınıf kendilerine kanıt kağıdı verildiği zaman aktif olarak argümantasyona katılmışlardır. Kanıt kağıdı verildiği zaman öğrencilerin yalnızca bilimin ürün aşamasına odaklandığı ve bilimin süreç boyutunu ihmal ettikleri düşünülmektedir. Çünkü kanıt kağıdında yer alan bilgiler bilimin ürünü olan içerik bilgisini temsil etmektedir. Bu nedenle sadece içerik bilgisine odaklanan öğrenciler doğru olan bilgiye ulaştıklarını düşündükleri zaman daha fazla tartışma yapmaya gerek duymamış ve dolayısı ile argümantasyon sürecine daha az dahil olmuş olabilirler (Kind vd., 2011).

Son araştırma sorusunda 6. sınıf öğrencilerinin argümantasyon sürecine katılımlarının doğası araştırılmıştır. Araştırma sürecine katılımın doğasının anlaşılması için Sampson ve Clark'ın (2011) ortaya koyduğu argümantasyon analizi boyutları (ör; karşıt görüş) kullanılmıştır. Gözlemlenen sınıfların argümantasyon sürecine katılımı ile ilgili ortaya çıkan benzerlikler ve farklılıklardan hareketle toplam sekiz iddia ortaya konmuştur. Bu iddialardan yedisi sınıflar arası benzerliklerden hareketle oluşturulmuş, bir iddia ise sınıflar arası farklılıklardan yola çıkılarak ortaya konmuştur. Benzerliklerden hareketle oluşturulan iddialardan üç tanesi uygulanan etkinlikler ile ilgilidir. Uygulanan etkinlikler ile ilgili iddialar şu şekildedir:

İddia 1. Öğrenciler argümantasyon sürecine en çok kendilerine kanıt kağıdı verildiği zaman katılmaktadır.

Bu iddia ile ilgili olarak daha önce yapılan çalışmalarda bu iddiaya benzer şekilde öğrencilerin kendi edindikleri verilerden çok hazır verileri kullandığını göstermektedir (Kolsto & Ratcliffe, 2007; Walsh & McGovan, 2017). Bu durumun nedeni hazır verinin direk kanıt olarak kullanılabilmesi ancak öğrencilerin kendi ürettikleri verinin henüz kanıta dönüşmemiş olması olabilir. Bu sebepten ötürü öğrenciler kendilerine kanıt kağıdı verildiği zaman argümantasyon sürecine daha fazla katılmış olabilirler. Ayrıca insanların kendi yaptıkları işten öte uzmanların ne dediğini önemsemesi de öğrencilerin kanıt kağıdındaki bilgileri aktif şekilde kullanma sebebi olabilir (Kolsto & Ratcliffe, 2007). Benzer şekilde sınıf tartışmalarının 40 dk ile sınırlı olması öğrencilerin kendi verileri yerine kanıt kağıtları üzerine odaklanmasına neden olmuş olabilir. Ayrıca, Kim ve Song (2006) öğrencilerin eksik bilgilerini saklamak için kendi düşüncelerini değil kanıt kağıdını kullandığını belirtmiştir. Konu bilgisi eksik olan öğrenciler bu çalışmada bu nedenle kanıt kağıdı verildiğinde etkin bir şekilde tartışmaya katılarak bu haftalarda sınıfta daha fazla tartışma yapılmasına neden olmuş olabilirler.

İddia 2. Öğrenciler ilerleyen haftalarda benzer etkinlikler yaptıkları zaman argümantasyon sürecine daha az katılmaktadır.

Öğrenciler karşılaştıkları etkinlikleri ilgi çekici bulmazlarsa, bu etkinliği gereksiz bulmakta ve bu etkinlikle ilgili soru sormamaktadır. Soru sorulmadıktan sonra da öğrenciler argümanlarını derinleştirememektedir (Chin & Osbourne, 2010). Bu çalışmada 2. haftada yapılan ısı yalıtımı etkinliğinin 1. haftada yapılan ısı iletimi etkinliğine benzemesi nedeniyle, 2. hafta etkinliği öğrencilerin ilgilsini çekmemiş olabilir ve bu yüzden öğrenciler argümantasyon sürecine çok fazla dahil olmamış olabilir. Isı yalıtımı haftasında benzer etkinlikler yapılmasındansa termos yapma gibi ilgi çekici etkinlikler yapılsaydı öğrenciler bu haftada argümantasyon sürecine daha fazla dahil olabilirdi.

İddia 3. Koşullar ideale yaklaştıkça öğrencilerin argümantasyon sürecine katılımları azalmaktadır.

Pilot çalışmada tel direnci ve telin özellikleri arasındaki ilişki araştırılırken öğrencilerin deneyi doğru yapamadığı gözlemlenmiştir. Bu nedenle, deney ana çalışmada ideal koşuların yer aldığı bir simülasyon programı ile değiştirilmiştir. Ana çalışmada ideal koşulların olduğu simülasyon programını takip eden öğrenciler doğru içerik bilgisine ulaşmalarına rağmen bu hafta çok fazla tartışma yapmamışlardır. Simülasyonda ideal koşulların hata içermemesi öğrencilerin farklı düşünmemesine neden olmuş olabilir. Benzer adımları takip edip, benzer sonuçlara ulaşan gruplar daha fazla tartışma yapma gereği hissetmemiş olabilirler. Bu noktada, Jimenez-Aleixandre (2008), öğretim programlarında kullanılacak etkinliklerin ikilemler içermesi gerektiğini belirtmektedir. Ayrıca, öğrenciler katıldıkları etkinliklerde farklı sonuçlara ulaşırsa daha fazla argümantasyon sürecine katılabilirler (Jimenez-Aleixandre, 2008).

Sınıflar arası benzerlikler ayrıca argümantasyon sürecine katılım boyutları ile ilgili dört iddia ortaya konmasını sağlamıştır. Buna göre;

İddia 1. Öğrenciler argümantasyon sürecinde en çok kendi fikirlerini açıklamaya odaklanmaktadır.

Bu çalışmada öğrencilerin başkalarının fikirlerini dinleyip eleştirmek, soru sormak, birlikte bilgi üretmek yerine keni fikirlerini ortaya koymak istedikleri görülmüştür.Benzer şekilde, Cavagnetto vd. (2010) öğrencilerin argüman üretmeye argüman eleştirmekten daha yatkın olduğunu rapor etmiştir. Öğrencilerin kendi fikirlerini ifade etmesi argümantasyon sürecini başlatması açısından önemli bir özellik olabilir. Bu çalışmada ayrıca öğrencilerin ilgisini çekmediği taktirde ilgili etkinliklerde fikir üretmedikleri gözlemlenmiştir. Öte yandan ön bilgi sahibi olmadıkları konularda ve kendilerine bu konularda kanıt kağıdı verildiğinde öğrenciler kendi argümanlarını oluşturmaya odaklanmaktadır. Fakat bu durumda argümantasyona katılımın diğer üç boyutunu (ör; karşıt fikir oluşturma) öğrenciler kullanmamaktadır.

İddia 2. Öğrenciler konu hakkında bilgi sahibi olduklarında ve kendilerine kanıt kağıdı verildiğinde karşıt görüşleri eleştirmeye odaklanmaktadır.

Karşıt görüşleri eleştirme daha önceki çalışmalarda argümantasyonun temel özelliği olarak belirtilmiştir (Sampson & Clark, 2011). Bu çalışmada da bu boyut 2. en sık kullanılan boyut olarak bu fikri desteklemektedir. Ayrıca bu çalışma öğrencilere kanıt kağıdı verilirse ve konu hakkında öğrencilerin ön bilgisi varsa öğrencilerin sıklıkla karşıt fikirleri eleştirdiğini göstermektedir. Bu durum, Naylor vd.'nin (2007) açıklamaları ile de uyumludur. Buna göre Naylar vd. (2007) öğrencilerin konuyu bildiklerinde daha fazla tartışma eğiliminde olduğunu ortaya koymuştur.

İddia 3. Öğrenciler deney yaptıkları haftalarda deney yapmadıkları haftalara oranla daha fazla soru sorarak bilgi aramaktadır. Öğrencilerin soru sorarak bilgi araması eğer deneyler ünite başında yapılırsa en yüksek seviyeye çıkmaktadır.

Bu iddiaya göre öğrenciler kendileri için yeni olan ve fazla bilgi sahibi olmadıkları konularda soru sorarak bilgi aramaktadır. Soru sorulma sebepleri genelde öğrencilerin bilgilerini artırmak istemesi, tutarsızlıkları ortaya koymak istemesi ve gözlenen durumların teorik açıklamalarını öğrenmek istemesidir. Benzer şekilde, Chin ve Osbourne (2010) tutarsızlıkla karşılaşan kişilerin sorular sorduğunu ve bu soruların da argümanları daha somut hale getirdiğini ifade etmektedir. Kanıt kağıdı varken öğrencilerin soru sormaması da belirsizlik ve bilgisizlik durumunda soru sorulduğu düşüncesini desteklemektedir çünkü kanıt kağıtlarında yer alan bilgiler belirsizliği ve bilgisizliği ortadan kaldırmaktadır.

İddia 4. Öğrenciler argümantasyon sürecinde genellikle bilgiyi birlikte oluşturma yolunu tercih etmemektedir.

Bu çalışmmada bilgiyi birlikte oluşturma yolu çok fazla kullanılmasa da bazı durumlarda öğrencilerin bilgiyi beraber oluşturmaya çalıştığı gözlemlenmiştir. Buna göre bir grup zorluk çektiğinde diğer gruplar o gruba fikir vererek yardımcı olmaktadır. Benzer şekilde, gruplar ortak sonuca ulaştığında birbirlerini desteklemek için beraber bilgi oluşturma yoluna gitmektedir. Ayrıca, gözlemlenen durumun teorik açıklamasını yapmakta eksik kaldıklarında yani öğrencilerin bilgisi yetersiz olduğunda öğrenciler birlikte bilgi oluşturma yolunu tercih etmektedir. Bu durum, Naylor vd. (2007) tarafından ortaya konan argümantasyon sadece zıt fikirlerin karşılaştırılmasıyla değil aynı zamanda benzer fikirlerin birleştirilmesiyle yapılır tezini desteklemektedir.

Sınıfların argümantasyon sürecine katılımları ile ilgili ortaya çıkan farklılıklara bağlı olarak ise bir iddia ortaya atılmıştır. Bu iddia şu şekildedir:

İddia. Genel olarak argümantasyon sürecine daha çok katılan sınıfta yer alan öğrenciler deneylerin yapıldığı haftalarda argümantasyon sürecine daha çok katılırken, genel olarak argümantasyon sürecine daha az katılan sınıfta yer alan öğrenciler deneylerin yapılmadığı haftalarda argümantasyon sürecine daha çok katılmaktadır.

Bu çalışmada deney yapılan hafta sayısının deney yapılmayan hafta sayısına göre daha fazla olması deney yapıldığında sürece daha fazla katılan sınıfın toplamda sürece daha fazla katılmasını sağlamış dolayısı ile bu sınıf daha aktif sınıf olarak isimlendirilmiştir. Öte yandan, daha aktif olan sınıf doğrudan veri kaynakları kullanıldığında sürece katılmakta ve dolaylı veri kaynağı kullanıldığında sürece daha az katılmaktadır. Bu durumun tersi ise daha az aktif olan sınıf için geçerlidir. Sonuç olarak hiç bir sınıf hem doğrudan hemde dolaylı olarak veri kaynaklarını kullanmamıştır. Wallace (2004) bu durumu kişilerin epistemolojik inançlarının eksikliği olarak görmektedir. Wallace'a (2004) göre epistemolojik inançları yüksek olan kişiler hem doğrudan kendi oluşturdukları verileri hemde dışarıdan dolaylı olarak edindikleri verileri argüman oluşumunda kullanırlar.

**Öneriler;**

Birinci olarak araştırmacılar epistemolojik inancı sadece nicel yollarla değil nitel yollarla da araştırmalıdır çünkü nicel verilerde gelişmiş epistemolojik inanca sahip olduğunu düşündüğümüz öğrencilerin sınıf içi gözlemlerde argümantasyon sürecinde yaptıkları şeyler onların gelişmemiş epistemolojik inanca sahip olduğunu göstermektedir. İkinci olarak, argümantasyon eğitimlerinde doğrudan açık bir şekilde epistemolojik inanç eğitimi verilmelidir. Bu şekilde öğrencilerin epistemolojik inançları gelişebilir. Ayrıca bu çalışmada öğrenciler kendilerine verilen kanıt kağıtlarının doğruluğunu güvenilirliliğini sorgulamamıştır. Bu nedenle argümantasyon çalışmalarında öğrencilere hatalı kanıt kağıtları da verilerek öğrencilerin bilginin doğruluğunu sorgulaması istenebilir.

Argümantasyon şeması ile ilgili olarak bu çalışmada ATBÖ eğitimi alan öğrencilerin zamanla sayıca daha fazla argümantasyon şeması kullanmadığı fakat daha fazla çeşit argümantasyon şeması kullandığı ve akıllı yürütme becerilerini artırdırğı gözlemlenmiştir. Bu nedenle akıl yürütme becerilerinin gelişimine odaklanmak isteyen araştırmacılar toplam kullanılan argümantasyon şeması sayısına değil kullanılan argümantasyon şeması çeşidine önem verebilirler. Ayrıca araştırmacılar ve öğretmenler bu çalışmada öğrencilerin sıklıkla kullanmış oldukları argümantasyon şemalarını göz önünde bulundurarak etkinlikler hazırlayabilirler. Benzer şekilde hazırlanan etkinlikler öğrencilerin hiç kullanmadıkları argümantasyon şemalarını da kullanmalarını sağlamalıdır. Öğrencilerin argümantasyon şemaları incelendiğinde uzman görüşü yoluyla argümantasyon şeması kullandıklarında hipotez yoluyla argümantasyon şeması kullanmadıkları, hipotez yoluyla argümantasyon şeması kullandıklarında ise uzman görüşü yoluyla argümantasyon şeması kullanmadıkları görülmüştür. Bu nedenle araştırmacıların ve öğretmenlerin öğrencilere hem doğrudan hemde dolaylı yoldan veri kullanımını teşvik etmeleri istenebilir.

Öğrenciler eğer bunu yaparlarsa gözlemleri ile teorik bilgiyi karşılaştırarak daha iyi öğrenme fırsatı yakalayabilirler.

Öğrencilerin argümantasyon sürecine katılımı ile ilgili olarak bu çalışmada öğrencilerin ilgi çekici olmayan etkinliklerde argümantasyon sürecine katılmadıkları tespit edilmiştir. Bu yüzden öğreniclerin ilgisini çekebilecek ve onları şaşırtarak motive edebilecek etkinlikler seçilirse öğrenciler argümantasyon sürecine daha fazla katılabilirler. Benzer şekilde, araştırmacılar ve öğretmenler ideal koşulların olduğu ve hata payının olmadığı etkinlikler kullanmamalıdır. Bu tarz etkinlikler öğrencilerin farklı şekilde veri toplamasını, farklı şekilde düşünmesini ve farklı sonuçlara ulaşmasını engellemektedir. Buda öğrencilerin argümantasyon sürecine katılımını olumsuz yönde etkilemektedir. Araştırmacılar ayrıca çalışma için neyi önemli görüyorsa önemli gördükleri bu şeyi çalışma boyunca sabit tutmalıdır. Örneğin bu çalışmada içerik bilgisi pilot çalışmada önemli görülürken, öğrencilerin argümantasyon sürecine katılımı ana çalışmada önemli görülmüştür. Bu durum başlangıçta fark yaratmıyor gibi görülsede durum öğrencilerin argümantasyon sürecine katılımı açısından fark yaratmaktadır. Buna göre; pilot çalışmada son haftada öğrenciler doğru içerik bilgisine ulaşamamış fakat argümantasyona sıklıkla katılmışlardır. Ancak temel amaç burada içerik bilgisi olduğu için deney, simülasyon programı ile değiştirilmiştir. Ana çalışmada yer alan simülasyon programı aracılığı ile öğrenciler doğru içerik bilgisine ulaşmış fakat bu seferde öğrenciler argümantasyon sürecine çok fazla katılmamıştır yani ana çalışmada önemli görülen temel amaç öğrencilerin argümantasyon sürecine katılımı bu durumdan olumsuz etkilenmiştir. Bu nedenle araştırmacılar argümantasyon çalışması yaptıkları zaman çalışma başında tek bir temel amaca odaklanmalı ve bu amacı çalışma boyunca değiştirmemelidir. Aksi taktirde bu çalışmanın son haftasında olduğu gibi sıkıntılar yaşanabilir. Argümantasyon sürecine katılım boyutları ile ilgili olarak bu çalışmada öğrencilerin en fazla kendi fikirlerini ortaya koyarak argüman üretme yolunu tercih ettikleri görülmüştür ve bu argümanlar argümantasyon sürecini başlatmaktadır. Bu nedenle daha sonra yapılacak olan çalışmalarda argümantasyon sürecini başlatan ve en çok kullanılan öğrencilerin kendi fikirleri ile argüman üretme boyutunu derinlemesine incelemeleri önerilmektedir çünkü bu boyutla ilgili çok çalışma yoktur ve bu boyut hakkında bilgimiz sınırlıdır. Ayrıca, öğrenciler konuyu bilmediklerinde ve kanıt kağıdı kullandıklarında sadece kendi fikirlerini öne sürmekte ve tartışmamaktadır, bu durumu engellemek için öğrenciler konuyu bilmediklerinde kanıt kağıdı vermek yerine ilgilerini çekecek soru sormalarını sağlayacak etkinlikler öğrencilere yaptırılabilir. Bu şekilde, öğrenciler argümantasyon sürecine daha fazla katılabilirler. Argümantasyon sürecinde karşıt fikir oluşturma ile ilgili olarak öğrenciler bildikleri konularda ve kendilerine kanıt kağıdı verildiğinde argümantasyon sürecine sıklıkla girmekte ve zıt fikirleri eleştirebilmektedir. Bu bilgi kısa süreli yapılacak argümantasyon çalışmalarında kullanılabilir. Normalde öğrencilerin argümantasyona karşı bir alışma süreci geçirmeleri gerekirken, kendi bildikleri konular kanıt kağıtları ile desteklenirse öğrenciler sürece daha kolay uyum sağlayabilirler. Bu durum araştırmacılar açısından zaman ve emek israfını engelleyebilir. Soru sorarak bilgi arama boyutu da bu çalışmada farklı fikirlerin ortaya çıkartılması açısından önemli bir argümantasyon sürecine katılım boyutu olarak görülmüştür. Soru sorarak bilgi arama boyutu argümantasyonu derinleştidiği için öğretmenler öğrencileri süreç boyunca birbirlerine soru sormaları konusunda teşvik edebilirler. Bu çalışmada ayrıca öğrencilerin ihtiyaç halinde, birbirlerini desteklemek için ve teorik bilgiyi bilmedikleri zaman birlikte bilgiyi oluşturmaya çalıştıkları gözlemlenmiştir. Bu durumlar göz önüne alınarak araştırmacılar ve öğretmenler etkinlikler hazırlarlarsa diğerlerine göre daha az kullanılan birlikte bilgi oluşturma süreci öğrenciler tarafından daha sonraki çalışmalarda daha fazla kullanılabilir. Bu çalışmada ayrıca öğrencilerin konu kazanımlarına ulaşması amaçlanmıştır. Bu nedenle öğrencilerin bazı araştırma soruları kazanımlarla uyumlu olmadığı için değiştirilmiş ve bu durum öğrencilerin çalışmaya katılımını olumsuz etkilemiştir. Bu yüzden öğretim programını yapan kişiler öğretmenlere daha fazla serbestlik verebilir ve konu bazında belirli kazanımlara odaklanmaktansa konu ile ilgili ana düşüncelere odaklanabilirler. Bu şekilde öğrencilerde konu ile ilgili istedikleri araştırma konusunu araştırabilirler. Ayrıca, bu çalışmada deney ve tartışma içeren konularda ATBÖ yaklaşımı uygulanmıştır. Öğretim programında diğer konularda da deney ve tartışma içeren kazanımlar eklenirse o konularda da araştırmacılar ATBÖ yaklaşımını uygulayabilirler. Bir diğer öneri bu çalışmanın nicel kısmının araştırma deseni ile ilgilidir. Bu çalışmada kontrol grubu kullanılmamıştır dolayısı ile öğrencilerin içerik bilgisinin gelişiminin tek nedeni ATBÖ olmayabilir. Bu nedenle benzer araştırmayı yapacak olan araştırmacıların kontrol grubu kullanması önerilmektedir. Nitel kısım ile ilgili olarak öğrencilerin argümantasyon sürecine katılımı incelenirken bu çalışmada sadece sınıf içi tartışmalara odaklanılmıştır. Daha sonra yapılacak çalışmaların hem konu öncesi yapılan tartışmalara, hem grup içi tartışmalara hemde sınıf içi tartışmalara odaklanması önerilmektedir. Bu şekilde argümantasyon sürecine katılımın doğası daha iyi aydınlatılabilir. Son olarak bu çalışma görece olarak kısa sürmüştür (6 hafta) ve daha uzun soluklu sürecek ATBÖ çalışmalarının ATBÖ ile ilgili daha detaylı bilgi vereceği düşünülmektedir.

## R. THESIS PERMISSION FORM / TEZ İZİN FORMU

**ENSTİTÜ /** INSTITUTE

**Fen Bilimleri Enstitüsü** / Graduate School of Natural and Applied Sciences ☐

**Sosyal Bilimler Enstitüsü** / Graduate School of Social Sciences ☒

**Uygulamalı Matematik Enstitüsü** / Graduate School of Applied Mathematics ☐

**Enformatik Enstitüsü** / Graduate School of Informatics ☐

**Deniz Bilimleri Enstitüsü** / Graduate School of Marine Sciences ☐

**YAZARIN /** AUTHOR

**Soyadı** / Surname : ŞEN
**Adı** / Name : Mehmet
**Bölümü** / Department : İlköğretim/ Elementary Education

**TEZİN ADI /** TITLE OF THE THESIS (**İngilizce** / English):
INVESTIGATING THE EFFECTIVENESS OF ARGUMENT-BASED INQUIRY ON 6$^{TH}$ GRADE STUDENTS' SCIENTIFIC LITERACY AND PORTRAYING THEIR ARGUMENTATION SCHEMES AND ENGAGEMENT IN ARGUMENTATION PROCESS

**TEZİN TÜRÜ /** DEGREE: **Yüksek Lisans** / Master ☐ **Doktora** / PhD ☒

1. **Tezin tamamı dünya çapında erişime açılacaktır. /** Release the entire work immediately for access worldwide. ☒

2. **Tez iki yıl süreyle erişime kapalı olacaktır.** / Secure the entire work for patent and/or proprietary purposes for a period of **two years**. * ☐

3. **Tez altı ay süreyle erişime kapalı olacaktır.** / Secure the entire work for period of **six months**. * ☐

***** *Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. / A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.*

**Yazarın imzası** / Signature ............................ **Tarih** / Date ............................